A. MEILLET
Professeur au Collège de France et à l'Ecole des Hautes Etudes
Membre de l'Institut

LA METHODE COMPARATIVE EN LINGUISTIQUE HISTORIQUE

PARIS
LIBRAIRIE HONORÉ CHAMPION, ÉDITEUR
7, QUAI MALAQUAIS (VI^e^)

1966

INSTITUTTET FOR SAMMENLIGNENDE KULTURFORSKNING

INSTITUTTET FOR
SAMMENLIGNENDE
KULTURFORSKNING

SERIE A: FORFLESNINGER

II.

A. MEILLET:

LA MÉTHODE COMPARATIVE EN LINGUISTIQUE HISTORIQUE

OSLO 1925
H. ASCHEHOUG & CO. (W. NYGAARD)

LEIPZIG
OTTO HARRASSOWITZ

PARIS
HONORÉ CHAMPION

LONDON
WILLIAMS & NORGATE, LTD

CAMBRIDGE, MASS.
HARVARD UNIVERSITY PRESS

INSTITUTTET
FOR SAMMENLIGNENDE KULTURFORSKNING

LA MÉTHODE COMPARATIVE EN LINGUISTIQUE HISTORIQUE

PAR

A. MEILLET

PARIS
LIBRAIRIE HONORÉ CHAMPION, ÉDITEUR
7, QUAI MALAQUAIS (VIe)

1966

© Editions CHAMPION Paris. Reproduction et traduction, même partielles, interdites.
Tous droits réservés pour tous pays, y compris l'U. R. S. S. et les pays scandinaves.

AVANT-PROPOS

Quand M. Alf Sommerfelt m'a annoncé l'intention qu'avait la grande institution nouvellement créée à Kristiania (maintenant Oslo), sur l'initiative de M. le professeur Fr. Stang, *Instituttet for Sammenlignende Kulturforskning,* de m'inviter à prendre part à la série de conférences qui devait marquer l'inauguration, j'ai aussitôt accepté avec joie.

Le jeune savant qui me faisait prévoir l'invitation avait travaillé durant plusieurs années — des années de la grande guerre — auprès des linguistes français. Il était devenu, pour tous les membres de notre petit groupe, un ami. J'étais heureux de le retrouver dans sa patrie où la linguistique lui doit tant, de pouvoir de nouveau, un peu de temps, travailler avec lui.

Et je me félicitais d'être associé aux débuts d'une institution dont les sciences de l'homme se promettent de grands résultats. On a beaucoup fait jusqu'ici avec les faits dont on dispose ; mais les matériaux qui se trouvent à la disposition des travailleurs ont été souvent exploités. Le moment est venu d'entreprendre de larges enquêtes systématiques qui fourniront un matériel nouveau et grâce auxquelles on pourra élargir et approfondir les théories actuellement admises. D'autre part, il importe de mettre en contact tous les savants qui s'occupent de l'homme et de ses civilisations à tous égards : une langue ne se comprend pas si l'on n'a pas une idée des conditions où vit la population qui l'emploie ; et l'on ne peut davantage comprendre vraiment une religion ou des usages sociaux sans connaître la langue des hommes qui pratiquent ces usages. L'unité du nouvel Institut exprime l'unité de l'objet qu'il étudie, à savoir l'homme. Il est à souhaiter que des organisations semblables soient fondées ailleurs et pourvues des larges ressources nécessaires aux enquêtes.

Du reste le sujet que, d'accord avec les organisateurs, j'ai choisi pour ces leçons est de ceux sur lesquels le moment est venu de réfléchir. Après une trentaine d'années où l'on a tiré des principes posés entre 1875 et 1880 les résultats qu'on en pouvait attendre, la linguistique historique est revenue à une période de fermentation. Des procédés nouveaux d'enquête ont apporté des résultats inattendus. Jamais on n'a fait pareil effort pour ramener à des « lois », à des « lois générales » même, tous les changements linguistiques, et jamais on n'a cherché à serrer d'aussi près les faits les plus particuliers, à pénétrer l'âme même des hommes chez qui se font les innovations. Des langues les plus anciennes où les changements ne nous apparaissent que réduits à des schémas jusqu'aux parlers actuels où les faits sont si concrets et si particuliers que le détail nous y dissimule les grandes tendances, on a observé des faits infiniment divers. La linguistique a pris contact avec toutes les disciplines voisines où l'on peut espérer trouver des explications. — Bien que trop lentement, l'étude historique des langues autres que les langues indo-européennes progresse ; mais les procédés qui ont réussi sur le domaine indo-européen ne sont pas partout également utilisables. Il faut réfléchir sur les méthodes employées, en examiner la légitimité, et voir comment on en pourrait étendre l'usage et les assouplir — sans en diminuer la rigueur — pour les conformer aux exigences des recherches sur des domaines nouveaux.

Il convient d'autant plus d'examiner les méthodes que, dans les derniers temps, beaucoup de linguistes ont avancé des hypothèses mal démontrées. Les nouvelles étymologies pullulent, et la plupart se présentent d'une manière telle qu'on n'en aperçoit pas même un commencement de démonstration. Il serait vain de les critiquer en détail tant qu'on n'est pas d'accord sur les conditions dont la réunion prouve la justesse d'un rapprochement étymologique. Pour qui admettra le bien fondé des principes exposés ici, une large part des hypothèses compliquées qui ont été faites sur le préindo-européen ou des étymologies indo-européennes qui ont été récemment avancées apparaîtra comme appelant à peine un examen. Bien que la discussion n'ait presque aucune place dans ces leçons, on y trouvera la critique implicite de beaucoup de travaux nouveaux qui ne satisfont pas aux exigences d'une méthode sévère.

On ne s'est pas proposé ici d'exposer des idées neuves, mais seulement de déterminer d'une manière précise les conditions où peut et doit s'employer en linguistique historique la méthode comparative. On s'estimera satisfait si le lecteur y trouve exactement marquées la valeur, mais aussi les limites de cette méthode.

Les idées développées ici n'ont été fixées par écrit qu'après avoir été exposées oralement. Rédigées après avoir été faites, ces conférences tiennent compte d'observations qui ont été présentées à l'auteur, en particulier par M. Sommerfelt que je suis heureux de remercier.

Je souhaite — mais il serait chimérique d'espérer — que ces exposés trouvent auprès du public universel un accueil aussi indulgent et aussi bienveillant que celui qu'ils ont reçu à Kristiania.

Kristiania (Oslo), septembre 1924.

A. M.

I.

DÉFINITION DE LA MÉTHODE COMPARATIVE.

Il y a deux manières différentes de pratiquer la comparaison : on peut comparer pour tirer de la comparaison soit des lois universelles soit des indications historiques. Ces deux types de comparaison, légitimes l'un et l'autre, diffèrent du tout au tout.

Il se rencontre partout dans le monde des contes d'animaux : les ressemblances entre les divers animaux et l'homme sont telles qu'il est naturel de prêter aux animaux des aventures comparables à celles qui arrivent aux hommes et d'exprimer par là des choses qu'il ne serait pas aussi facile de faire entendre directement. On peut comparer ces contes entre eux pour en définir les formes, les caractères, l'emploi, et pour faire ainsi une théorie générale des contes d'animaux. Les concordances que l'on constate résultent de l'unité générale de l'esprit humain, et les différences de la variété des types et des degrés de civilisation. On aboutit ainsi à s'instruire sur les caractères généraux de l'humanité, mais on n'apprend rien sur son histoire.

Si l'on examine, avec un jeune savant français, M. Dumézil, les mythes indo-européens relatifs à la boisson d'immortalité, on obtient des résultats tout autres. L'idée qu'il y aurait une boisson propre à conférer l'immortalité est trop naturelle pour être caractéristique. Mais, quand on rencontre, d'une manière plus ou moins complète, chez les divers peuples de langue indo-européenne, la légende d'une boisson d'immortalité fabriquée dans une cuve gigantesque, quand à cette légende se joint l'histoire d'une fausse fiancée, et le récit d'une lutte entre des dieux et des êtres démoniaques, il y a là un ensemble de faits singuliers qui n'ont pas en eux-mêmes de liens entre eux et dont la réunion ne saurait, par suite, être fortuite.

Si le sens à exprimer par la langue était lié par un lien naturel, lâche ou étroit, aux sons qui l'indiquent, c'est-à-dire si, par sa valeur propre, en dehors de la tradition, le signe linguistique évoquait en quelque manière une notion, le seul type de comparaison utilisable pour le linguiste serait le type général, et toute histoire des langues serait impossible.

Mais, en fait, le signe linguistique est arbitraire : il n'a de valeur qu'en vertu d'une tradition. Si l'on exprime en français l'unité par *un, une,* la dualité par *deux,* etc., ce n'est pas parce que les mots *un, une — deux —*, etc., ont par eux-mêmes un rapport quelconque avec l'unité, la dualité, etc., mais uniquement parce que tel est l'usage enseigné par ceux qui parlent à ceux qui apprennent à parler.

Seul, le caractère totalement arbitraire du signe rend possible la méthode comparative historique qui va être étudiée ici.

Soit les noms de nombre en français, en italien et en espagnol. On a une série :

FRANÇAIS	ITALIEN	ESPAGNOL
un, une	*uno, una*	*uno, una*
deux	*due*	*dos*
trois	*tre*	*tres*
quatre	*quattro*	*cuatro*
cinq	*cinque*	*cinco*
six	*sei*	*seis*
sept	*sette*	*siete*
huit	*otto*	*ocho*
neuf	*nuove*	*nueve*
dix	*dieci*	*dies*
vingt	*venti*	*veinte*
trente	*trenta*	*treinta*
quarante	*quaranta*	*cuarenta*
cent	*cento*	*ciento*

De pareilles concordances ne sauraient être accidentelles ; elles le sont d'autant moins que les différences d'une langue à l'autre se laissent ramener à des règles de correspondance définies : ainsi la différence entre *huit, otto* et *ocho* est grande au premier abord ; mais le fait qu'elle n'est pas accidentelle résulte de ce qu'il y a une série de correspondances semblables, ainsi fr. *nuit,* it. *notte,* esp. *noche,* ou fr. *cuit,* it. *cotto* ; et l'on a de même fr. *lait,* it. *latte,* esp. *leche* ; fr. *fait,* it. *fatto,* esp. *hecho* ; etc. Les concordances évidentes dès l'abord montrent la voie à suivre. Mais ce sont les règles de correspondances phonétiques qui seules permettent d'en tirer parti.

Là où les ressemblances visibles ont indiqué la bonne voie, il arrive souvent que tel détail *singulier* apporte une confirmation. Il est significatif par exemple qu'il y ait une distinction du masculin et du féminin pour *un, une* et pas pour les autres nombres.

On est donc conduit à poser que les noms de nombre du français, de l'italien et de l'espagnol remontent à une seule et même tradition originelle. En pareil cas, l'expérience montre qu'il y a deux types de tradition possibles : les trois groupes considérés peuvent remonter à une origine commune, ou bien deux des trois peuvent avoir emprunté les formes de l'autre. En l'espèce, la seconde hypothèse est exclue, parce qu'on ne saurait expliquer les formes d'aucune des trois langues par celles d'une autre. Ni le fr. *huit* ne peut sortir de it. *otto* ou de esp. *ocho,* ni it. *otto* de fr. *huit* ou de esp. *ocho,* ni esp. *ocho* de fr. *huit* ou de it. *otto*. Il est prouvé ainsi que les noms de nombre du français, de l'italien et de l'espagnol ont un point de départ commun qui n'est ni français, ni italien, ni espagnol.

Dans l'exemple choisi, les concordances sont si nombreuses, si complètes et les règles de correspondances si faciles à reconnaître, qu'elles sont propres à frapper immédiatement des profanes et qu'il n'y a pas besoin d'être linguiste pour en apercevoir la valeur probante. Les concordances sont moins frappantes et les règles de correspondances plus difficiles à déterminer si l'on observe des langues séparées par de plus grands intervalles dans l'espace et dans le temps, comme le sanskrit, le grec attique ancien, le latin et l'arménien classique :

	SKR.	GR. ATT.
	—	—
« un »	*ékaḥ, ékā, ékam*	*hēs, mia, hen*

LAT.	ARM.
ūnus, ūna, ūnum	*mi*

[là où il y a trois formes, l'une est celle du masculin, la seconde celle du féminin, la troisième celle du neutre ; l'arménien n'a pas de différences de genre grammatical.]

	SKR.	GR.	LAT.	ARM.
« deux »	*d(u)vā*	*dyo*	*duo*	*erku*

[on ne note ici que les formes du masculin, là où le genre est distingué ; de même pour « trois » et pour « quatre ».]

	SKR.	GR.	LAT.	ARM.
« trois »	*tráyaḥ*	*trēs*	*trēs*	*erekʿ*
« quatre »	*catvā́raḥ*	*téttares*	*quattuor*	*č̣orkʿ*
« cinq »	*páñca*	*pénte*	*quinque*	*hing*
« six »	*ṣáṭ*	*heks*	*sex*	*vec̣*
« sept »	*saptá*	*heptá*	*septem*	*ewtʿn*
« huit »	*áṣṭā*[1]	*óktō*	*octō*	*utʿn*
« neuf »	*náva*	*ennéa*	*nouem*	*inn*
« dix »	*dáça*	*déka*	*decem*	*tasn*

Si, réserve faite du nom de nombre « un », les correspondances entre le grec, le latin et même le sanskrit sont évidentes dans une large mesure, il n'en va pas de même de celles entre l'arménien et les autres langues.

---

[1] Forme védique devant consonne, alternant avec *áṣṭāv* devant voyelle.

Mais il suffit d'examiner les faits arméniens de près pour que la valeur probante des concordances ressorte.

Ainsi arm. *erku* « deux » ne ressemble pas à lat. *duo,* etc. ; mais d'autres correspondances montrent que *erk-* peut répondre à **dw-* d'autres langues ; ainsi, de même que le grec a pour l'idée de « craindre » une racine *dwi-,* l'arménien a *erki-* (*erkiwł* « crainte »), et de même que le grec a pour dire « longtemps » un vieil adjectif *dwārón,* l'arménien a *erkar* « long » (v. ci-dessous, p. 31. La concordance se laisse donc ramener à une règle générale de correspondance : un ancien *dw-* aboutit à arm. *erk-*).

Au premier terme des composés, le grec a *dwi-,* et l'arménien *erki-*. Il y a donc un groupe de concordances singulières qui ne laisse aucun doute (voir ci-dessous, p. 107).

Les formes arméniennes *erekʻ* et *c̣orkʻ* sont loin de gr. *trēs, téttares* ; mais elles se laissent, au moins en partie, expliquer par des correspondances semblables. Et, détail caractéristique, de même qu'en sanskrit et en grec, « trois » et « quatre » ont des formes casuelles d'un type ordinaire, les noms à partir de « cinq » sont invariables ; or, en arménien, « trois » et « quatre » ont des formes casuelles normales, et, en particulier, le *-kʻ* final est la marque du nominatif pluriel arménien, marque qui ne se retrouve pas aux autres cas.

Moins apparentes au premier coup d'œil que les concordances entre le français, l'italien et l'espagnol, les concordances des formes des noms de nombre en sanskrit, en grec, en latin et en arménien ne sont au fond pas moins certaines, on le voit.

Ces concordances, qui ne peuvent s'expliquer par des emprunts d'une langue à l'autre, supposent une origine commune. Mais il reste à les interpréter d'une manière systématique : tel est l'objet de la linguistique historique comparative.

Le procédé dont on vient de voir le principe peut sembler compliqué, difficile à manier. Mais il n'y en a aucun autre pour faire l'histoire des langues.

Car l'histoire des langues ne se fait jamais au moyen d'une suite de textes rangés en ordre chronologique. Si le linguiste se sert de textes anciens, ce n'est que pour y observer des états de langue. Il va de soi que, pour toutes les langues anciennes, les faits se laissent observer seulement à l'aide des textes. C'est sur des documents écrits qu'on observe l'attique ou le gotique, l'arménien classique ou le vieux slave. Interprétés avec critique, ces documents donnent beaucoup, et l'on peut souvent avoir une notion précise de certains états de langue anciens. Mais cette étude permet seulement de déterminer l'état d'une langue à un certain moment, dans certaines conditions. L'examen des textes n'est qu'un substitut de l'observation directe devenue impossible.

Même dans les meilleurs cas, la langue écrite est bien loin d'enregistrer exactement les changements successifs de la langue parlée. Souvent la langue écrite est fixée, et la forme qu'elle présente ne change presque pas d'un siècle à l'autre. Là même où elle n'est pas entièrement fixée, l'usage écrit est ordinairement dominé, dans une large mesure, par des formes antérieures — qui ne sont pas toujours connues.

Soit l'exemple classique du latin. Entre la langue écrite telle qu'on la trouve chez Plaute et celle qu'on trouve chez saint Augustin, il y a des différences de détail ; mais la graphie ne varie presque pas ; les formes grammaticales et le principal du vocabulaire demeurent les mêmes. Les philologues se sont donné beaucoup de mal pour trouver des différences entre les périodes du latin à date historique. Ils ont apporté à ces recherches toute la minutieuse précision qu'y peut mettre un grammairien. Et ils ont constaté en effet de menues différences ; mais, pour une assez large part, ces différences tiennent au genre littéraire : la comédie de Plaute, destinée au gros public, n'est pas comparable aux discours de Cicéron ni aux bulletins du puriste qu'était César. La langue de la poésie hellénisante n'est pas celle de la prose. Il y a aussi des différences d'un écrivain à l'autre. Mais, au fond, il n'a jamais été écrit et enseigné à l'école qu'un seul latin.

Du reste, entre les formes que livrent les textes écrits, seule la comparaison indique celles qui ont une valeur pour l'histoire ultérieure de la langue. Pour désigner l' « oreille », on trouve dans les textes écrits du latin *auris*, et aussi le dérivé *auricula*. Rien, dans les anciens textes latins, n'avertit que l'une ou l'autre de ces formes doive prévaloir sur l'autre ; *auris* est la forme courante. Or, c'est sur *auricula* que reposent les formes des langues romanes : fr. *oreille*, it. *orecchia*, esp. *oreja*. Seule, la comparaison du français, de l'italien, de l'espagnol, etc. renseigne sur la forme qui sert de base aux langues romanes.

D'autre part, la langue qui a survécu n'est pas celle qui s'écrivait. Entre le latin littéraire, qui est conservé par les textes des écrivains, et le latin parlé, que continuent les langues romanes, il y avait des différences, variables suivant les individus et suivant leur degré de culture. Or, les langues romanes ne continuent pas le latin littéraire. La « bouche » se nommait en latin écrit *ōs* ; mais ce qui a subsisté, c'est le nom vulgaire *bucca.* Et ainsi dans une foule de cas.

Là même où des textes d'époques diverses fournissent des états de langue successifs, on n'observe pas une continuité. Les changements essentiels auxquels est dû le passage du type latin ancien au type roman ont eu lieu entre le latin écrit et les premiers monuments des langues romanes. Sans doute les gens qui ont écrit le latin du IIIe au Xe siècle ap. J. C. ont laissé échapper bien des formes qui ne sont pas correctes en latin ancien et qui sont dues à la langue déjà toute différente qu'ils parlaient ; mais on ne peut apprécier la valeur de ces lapsus et y reconnaître des témoignages linguistiques que grâce à la connaissance des langues romanes. Les textes latins écrits depuis l'époque impériale jusqu'à l'époque carolingienne par des gens assez peu lettrés pour laisser échapper des fautes contre l'usage classique apportent des précisions, notamment dans la chronologie des faits, à la grammaire comparée des langues romanes ; mais, si les lumières qu'ils projettent sont précieuses, c'est pour éclairer la comparaison ; la comparaison venant à manquer, elles se perdraient dans le vide. Si les langues romanes étaient mortes avant d'avoir été notées, on apercevrait que le latin classique se corrompait ; mais on ne pourrait faire aucune théorie précise et complète des changements accomplis. Tout l'essentiel des changements linguistiques a lieu hors de notre vue.

L'histoire du perse fournit un exemple plus remarquable encore. On connaît — d'une manière très incomplète, il est vrai — le perse à l'époque de Darius et de

Xerxès, grâce aux inscriptions monumentales de ces rois. Puis les témoignages manquent du v^e siècle av. J.-C. jusqu'au III^e siècle ap. J.-C. Alors on trouve le « pehlvi. »

Mais, d'abord, ce pehlvi n'est pas, comme on l'a cru jusqu'aux dernières années, la continuation du parler qui a été noté dans les inscriptions des souverains achéménides : ainsi que je l'ai indiqué moi-même et que M. Tedesco l'a démontré d'une manière plus précise et plus complète, le pehlvi des textes de l'époque sassanide, ancêtre du persan littéraire, présente quelques traits qui le distinguent du perse achéménide. On ne saurait donc dire que le linguiste possède ici la tradition d'une même langue ; il observe, à des dates différentes, deux parlers de types très voisins, mais non identiques.

En second lieu, entre le perse achéménide et le pehlvi sassanide, il y a une différence de développement aussi grande que celle qui sépare le latin ancien du roman. Entre les deux dates, la langue a changé de caractère. Comme sur le domaine latin, l'essentiel du changement s'est passé hors de la vue de l'historien. Le linguiste dispose de deux états de langue profondément distincts l'un de l'autre ; pour faire une histoire, il lui faut restituer l'entre-deux. La comparaison des divers parlers iraniens modernes aide à faire cette restitution ; mais elle ne fournit que des directions, non des témoignages positifs.

D'une manière générale, l'histoire des langues ne se fait donc qu'en comparant des états de langue les uns aux autres. Car les faits que fournit la succession des textes dans les cas exceptionnels où les gens qui écrivent ont suivi plus ou moins complètement l'usage de la langue parlée en leur temps sont presque toujours d'importance médiocre, le plus souvent insignifiants, par rapport aux grands faits qui ont eu lieu sans avoir été notés par personne. Pour déterminer les états de langue du passé, le linguiste doit se servir de la philologie la plus exacte, la plus précise : et chaque progrès dans la précision philologique permet un progrès nouveau pour le linguiste. Le contact de plus en plus étroit qui s'est heureusement établi entre philologues et comparatistes est nécessaire pour que le linguiste puisse utiliser tous les faits, des faits certains, et des faits observés avec la dernière précision. Mais, à elle seule, la philologie n'apporte même pas un commencement d'histoire linguistique.

La comparaison est le seul instrument efficace dont dispose le linguiste pour faire l'histoire des langues. On observe les résultats des changements, non les changements eux-mêmes. C'est donc seulement à l'aide de combinaisons qu'on suit — et qu'on peut suivre — le développement des langues.

Mais ces combinaisons sont, on le verra, rigoureuses et précises. Toutes reposent sur l'affirmation que certaines concordances données entre langues diverses ne s'expliquent pas par des traits communs à tous les hommes et requièrent l'hypothèse d'une tradition particulière.

Telle est l'essence de la méthode comparative. Pour apprécier la valeur probante d'une combinaison, il n'y a qu'à ne jamais perdre de vue ce caractère de la preuve.

II.

LES LANGUES COMMUNES.

De ce que la méthode comparative est la seule qui permette de faire l'histoire des langues il résulte que, tant qu'une langue est isolée, elle est dénuée d'histoire. Entre l'état du basque au XVI^{e} siècle et l'état du basque aujourd'hui, il y a des différences ; mais les changements ne sont pas essentiels ; en substance, la langue est restée la même. Si donc on ne trouvait pas de moyen de rapprocher le basque de telle ou telle autre langue, il n'y aurait aucun espoir d'en faire jamais l'histoire. En revanche, si les essais faits — par M. Marr et M. Oštir, d'une part, par M. Trombetti, de l'autre, — pour rapprocher le basque d'un grand groupe de langues du bassin de la Méditerranée, et spécialement des langues caucasiques, aboutissent, le basque, sortant de son isolement, entrera dans l'histoire.

Pour tous les groupes actuellement établis et étudiés d'une manière méthodique, le moyen de faire le rapprochement est de poser une « langue commune » initiale (*Ursprache,* comme on dit en allemand). Ce n'est rien que de poser des rapprochements partiels : chaque fait linguistique fait partie d'un ensemble où tout se tient. Il ne faut pas rapprocher un fait de détail d'un autre fait de détail, mais un système linguistique d'un autre système. Sans doute il n'est pas toujours possible de restituer ainsi l'ensemble d'une langue par des procédés comparatifs ; on ne saurait même affirmer a priori qu'il faille, dans tous les cas, restituer une langue initiale unique — le problème des langues « mixtes » sera envisagé par la suite. Mais, là où elle aboutit pleinement, la comparaison aboutit à restituer une langue initiale.

Que vaut cette restitution ? Il n'est presque jamais possible de confronter la restitution avec une réalité connue. Mais il y a un cas où on le peut ; c'est celui des langues romanes. Or, la langue commune à laquelle on est conduit par la comparaison des langues romanes ne fournit pas — tant s'en faut — tout ce qu'était le latin au moment où se sont séparées les unes des autres les langues qui continuent le latin. Si l'on ne savait du latin que ce qu'enseignent les langues romanes, on ignorerait par exemple l'ancien futur du type *amabo* ou du type *dicam, dices*. Et surtout on n'aurait aucune idée de la déclinaison : jusqu'au xiii^e^ siècle, les parlers gallo-romans distinguent, dans le masculin, un cas sujet d'un cas régime ; les autres parlers romans n'offrent même pas cette distinction. Le substantif roman est, à chaque nombre, invariable partout dès la fin du xiii^e^ siècle. Or, à la date où s'est brisé l'Empire romain, la déclinaison subsistait et jouait encore un grand rôle. Si la comparaison fournit, à bien des égards, des données qui concordent avec la réalité attestée — ainsi pour la flexion verbale —, elle est donc loin de fournir le tout de la langue. — Certaines survivances, curieuses pour qui possède la forme ancienne, seraient inintelligibles pour le linguiste qui disposerait seulement de la comparaison : la nasale de fr. *rien* apparaît comme un reste de l'ancien accusatif quand on sait que l'accusatif latin était caractérisé par la désinence *-m* ; mais le romaniste qui utiliserait seulement la comparaison n'aurait aucun moyen d'y reconnaître un accusatif tel qu'était lat. *rem*. — La « restitution » fournit donc de la « langue commune » une idée incomplète, et sans doute très incomplète, le plus souvent.

C'est que les langues qui continuent une même « langue commune » n'en conservent pas seulement certains traits anciens. Longtemps, elles gardent une tendance à présenter des innovations ou identiques ou semblables, de telle sorte que certaines parties de la « langue commune » disparaissent partout sans laisser de traces, ou n'en laissent que d'impossibles à discerner si l'on ne connaît pas en fait la « langue commune ».

Dans toutes les langues romanes, la déclinaison des noms, qui était un élément si essentiel du latin ancien, s'est éliminée de bonne heure. L'élimination a eu lieu de manière indépendante dans chaque langue ; les parlers gallo-romans en portent témoignage puisque, après avoir conservé un reste de déclinaison sous forme d'opposition entre un cas sujet et un cas régime, ils ont à leur tour perdu dès la fin du XIII^e^ siècle cette survivance.

Partout dans les langues romanes, la nasale finale de l'ancien accusatif s'est amuie : le français n'a pas trace de la nasale finale de *terram* ou de *regem* dans ses formes *terre, roi*. Cette nasale ne s'est maintenue que dans le monosyllabe accentué *rem* qui a abouti à fr. *rien*.

Mais, si des innovations parallèles, graves et nombreuses, empêchent la restitution des comparatistes d'être complète, à beaucoup près, il y a, au fond de la restitution, une justesse. Ainsi la constance avec laquelle *-m* finale est notée dans l'orthographe latine classique ne donne pas une notion exacte de la réalité. Cette nasale finale se prononçait faiblement. Dans le latin le plus ancien, la graphie ne la note souvent pas. Pour les poètes, une *-m* finale n'empêche pas l'élision devant voyelle : en prosodie latine, *-am* de *terram* s'élide devant voyelle exactement comme *-a* de *terra*. L'absence de *-m* finale dans les langues romanes exprime la faiblesse de cette nasale en latin, faiblesse que la graphie classique du latin dissimule.

Du reste, pour qui se propose d'étudier les langues romanes, peu importent les traits du latin qui ont disparu sans retour. Ce qui pour lui est utile, ce sont les éléments qui ont servi à constituer les formes nouvelles prises par le latin. La restitution ne fournit pas le latin réel, tel qu'il se parlait ; et aucune restitution ne saurait fournir la « langue commune » telle qu'elle a été parlée. Ç'a été une hardiesse de génie, pour Schleicher, de « restituer » l'indo-européen à l'aide des langues de la famille historiquement attestées ; mais ç'a été, de sa part, une erreur grave que de composer un texte dans cette langue restituée. La comparaison apporte un système de rapprochements sur lequel on peut fonder l'histoire d'une famille de langues ; elle ne fournit pas une langue réelle, avec tous les moyens d'expression que comportait cette langue.

Entre la « langue commune » initiale restituée par comparaison et la langue attestée en fait, il peut s'intercaler une ou plusieurs « langues communes » intermédiaires. C'est ainsi que, entre l'indo-européen, d'une part, et les langues romanes, de l'autre, s'insère une grande « langue commune », le « roman commun », qu'on est convenu de nommer « latin vulgaire ». De même, entre l'indo-européen, d'une part, le gotique, le vieux haut allemand et le vieil anglais, le vieux norrois, de l'autre, il y a eu une « langue commune », le « germanique commun », langue non attestée en fait, mais dont l'existence est supposée par un ensemble d'innovations systématiques. Ces paliers facilitent beaucoup l'explication.

Si, pour expliquer les faits français, italiens, espagnols, portugais, roumains, les romanistes devaient remonter à l'indo-européen commun et ne possédaient pas le palier « roman commun », l'explication des faits resterait singulièrement incomplète, impossible le plus souvent.

Quand on ne dispose pas de pareils paliers, on rencontre de graves difficultés. C'est ce qui arrive aux romanistes quand ils sont en présence de parlers locaux gallo-romans ou italiens. Il n'y a pas eu de « gallo-roman commun », d' « italien commun ». On est donc réduit à comparer chaque parler gallo-roman ou italien au roman commun, sans intermédiaire. Ce n'est pas à dire que les parlers de la France du Nord, par exemple, n'aient pas entre eux beaucoup de traits communs dont les plus caractéristiques se retrouvent dans le français littéraire fondé sur le parler de Paris. Mais, depuis l'époque de séparation des parlers romans, chacun a son histoire propre. Les romanistes qui s'occupent des parlers gallo-romans et des parlers italiens se trouvent ainsi obligés soit à refaire pour chaque parler l'histoire entière du type auquel appartient le parler étudié, soit à employer des simplifications qui sont dans une certaine mesure arbitraires. Pour sortir de cette difficulté, il a été fait divers essais, notamment celui de M. Oscar Bloch sur les parlers d'une région des Vosges (celle de Remiremont). Quelque ingéniosité qu'on déploie pour éluder la difficulté, il n'est pas possible de s'en débarrasser tout à fait.

Quand on dispose de langues communes intermédiaires, l'explication des faits est beaucoup facilitée. Mais il n'est pas toujours possible d'en tirer complètement parti. Ainsi entre l'indo-européen et le latin, on sait qu'il y a eu deux périodes d'unité : une période italo-celtique et une période italique (l' « italique » ayant fourni le latin et l'osco-ombrien). Mais la période de communauté italo-celtique rend peu de services, parce qu'elle a été de trop courte durée pour entraîner des innovations nombreuses et profondes. Et la période de communauté italique n'en rend guère plus parce que l'osco-ombrien est connu de manière toute fragmentaire.

Pour chaque langue, le problème de la restitution de la « langue commune » initiale se pose d'une manière particulière. Il faut, en chaque cas, tirer parti des situations particulières qui se présentent.

« Langue commune » suppose civilisation commune. Car il n'y a en fait langue commune que là où un parler s'étend sur des domaines où il ne s'employait pas antérieurement. Il peut arriver que cette langue commune apportée par des conquérants ou des colons ne soit jamais acceptée par l'ancienne population indigène et que celle-ci émigre ou disparaisse : tel est en gros le cas de l'anglais dans l'Amérique du Nord. Il peut arriver aussi que les hommes qui apportent la langue nouvelle se croisent avec les indigènes et que, malgré le nombre minime des envahisseurs, les indigènes acceptent la langue nouvelle : tel est le cas de l'espagnol ou du portugais dans certaines parties de l'Amérique centrale et méridionale. Dans tous les cas, une langue ne s'étend que si elle est l'organe d'une civilisation douée de prestige. Et il arrive même que l'extension de la langue soit due tout entière au prestige d'une civilisation : la *koinè* ionienne-attique a remplacé tous les autres parlers grecs, parce qu'elle était l'organe de la civilisation hellénique par excellence.

Si le latin a été accepté dans toute la partie occidentale de l'Empire romain, c'est qu'il portait avec lui une civilisation supérieure à celle des peuples soumis par Rome. Dans la partie orientale de l'Empire, où le grec servait à une civilisation plus ancienne et, au point de vue intellectuel au moins, supérieure, le latin n'a pas gagné. En Gaule, l'aristocratie s'est mise à étudier le latin dès après la conquête, et dès les débuts de l'époque impériale, elle était de culture toute latine. Le gaulois n'a pas disparu pour cela ; le peuple en a longtemps encore gardé l'usage. Les artisans de la fabrique gauloise de poterie récemment explorée dans le Sud de la France se servaient encore du gaulois ; la grammaire qu'ils employaient était gauloise ; les noms de nombre étaient gaulois ; mais tout ce qui était de civilisation n'était déjà plus gaulois : les noms propres d'hommes ne sont pas gaulois ; les noms de vases fabriqués ne le sont pas non plus (voir à ce sujet les remarques de M. J. Vendryes, *Bulletin de la Société de linguistique*, XXV, p. 39 et suiv.). C'est le prestige d'une civilisation supérieure qui amène une population à changer de langue.

Cette supériorité n'est pas nécessairement d'ordre matériel. Elle peut consister en une organisation sociale particulièrement adaptée aux besoins d'une époque et d'une région données. Il n'y a pas de raison de croire que la nation dont l' « indo-européen commun » était la langue ait disposé de moyens matériels supérieurs à ceux de ses voisins, qu'elle ait par exemple été plus avancée pour l'agriculture ou pour le travail du métal. Ce qui caractérise les anciens peuples de langue indo-européenne, c'est le sens d'organisation sociale et la puissance d'initiative de leur aristocratie. Cette organisation ne comportait pas de pouvoir central. Chaque chef de famille était maître dans sa grande famille. Chaque chef de tribu était indépendant des autres. Tout au plus y avait-il des confédérations provisoires et instables. Partout où l'on observe les anciens peuples de langue indo-européenne, on voit des tribus autonomes et des familles où l'autorité est exercée par un *pater familias*. Chaque cité grecque est un petit État indépendant, de même que chaque groupe gaulois ou germanique. Dès qu'un chef se sent capable de grouper autour de lui des hommes entreprenants, il part pour quelque région où il puisse trouver le moyen de mener sa vie propre et autonome. La colonisation des rives de la Méditerranée par les anciens Grecs comme celle de l'Islande, de l'Angleterre, de la Normandie et jusqu'à la Sicile par les Vikings — S. Bugge, avec son intuition aiguë de la réalité, a déjà rapproché les deux mouvements — sont des exemples illustres de cet esprit d'indépendance et d'aventure. La nation indo-européenne se composait ainsi de petits groupes qui ont dû garder longtemps le sens de leur unité, et qui, même après le temps où l'unité générale du monde indo-européen était brisée, ont formé des nations du même type ayant aussi leur unité. Par exemple le groupe indo-iranien dont le nom indigène est *ărya-*, s'est étendu largement sur l'Asie : on l'aperçoit au XIIIe siècle en Cappadoce, et les Assyro-Babyloniens ont pris contact avec lui en Médie. La nation « aryenne » a poussé ses conquêtes très avant sur le plateau iranien, dans l'Inde, sur toute la rive méridionale de la mer Noire, et à l'Est jusqu'aux frontières de la Chine, où datant du début de l'ère chrétienne, il a été trouvé des textes sogdiens près de la grande muraille.

A se disperser ainsi et à porter des extrémités de l'Asie aux extrémités de l'Europe leur type d'organisation sociale avec leur langue, les peuples de langue indo-européenne ont perdu le sens de leur ancienne unité nationale et leur unité linguistique. Les nouveaux groupes qui se sont constitués sur chacun des domaines occupés ont perdu à leur tour leur unité. Le monde aryen, le monde germanique, le monde celtique se sont disloqués comme s'était disloqué le monde indo-européen. On est parvenu de cette manière au monde moderne où presque chaque pays a sa « langue commune », et où cette langue commune tend à effacer les parlers locaux. Il y a là un état de choses nouveau et qui n'est pas susceptible de durer à la longue : la multiplication des « langues communes » dans l'Europe d'aujourd'hui, en un temps où il y a au fond unité de civilisation matérielle et intellectuelle, est une anomalie.

Chacune des grandes « langues communes » du passé doit exprimer un type de civilisation. Et c'est pour cela que la plupart des langues du monde paraissent remonter à un nombre restreint de langues communes. Tant que les diverses grammaires comparées ne seront pas constituées, on ne pourra rien affirmer absolument. Mais, dès maintenant, on a notamment l'impression que toutes les langues nègres de l'Afrique reposent sur une même langue originelle. Il serait risqué de mettre les aires linguistiques en rapport exact, constant, régulier avec des aires de civilisation. Mais un rapport est certain. Et ce sera l'une des tâches de l'étude de l'homme dans l'avenir que de relier les langues communes aux aires de civilisation. Dès maintenant on entrevoit des concordances caractéristiques.

III.

DES PREUVES EMPLOYÉES.

Pour établir l'existence d'une ancienne langue commune, il faut retrouver dans les langues comparées les traits spécifiques de cette langue pour autant qu'ils se sont maintenus. Il y a donc à chercher comment se comportent les divers éléments de la langue; car ils ne se conservent pas également ni de la même manière. Toute langue comprend trois systèmes distincts qui sont liés les uns aux autres de certaines manières, mais qui sont susceptibles de varier en une large mesure indépendamment les uns des autres : la morphologie, le phonétisme et le vocabulaire.

La morphologie, c'est-à-dire l'ensemble des procédés par lesquels on modifie et on groupe les mots pour constituer des phrases, est ce qu'il y a de plus stable dans la langue. Mais il faut distinguer ici entre les procédés généraux et le détail des formes.

A les considérer d'une manière générale, les types morphologiques sont peu variés. La principale différence consiste en ceci que, dans certaines langues, les rapports entre les mots et les divers emplois des mots sont marqués par des formes spéciales qu'affectent les mots eux-mêmes, tandis que, dans d'autres langues, les rapports et les emplois sont marqués par des éléments additionnels, par des particules ou mots accessoires et par l'ordre des mots. Dans les premières langues, il y a, comme on dit, flexion. Cette distinction n'est du reste pas absolue, et il y a des mélanges des divers procédés à doses diverses.

Bien que l'usage fait de tel ou tel type se maintienne souvent très longtemps et laisse des traces alors même que le type tend à s'abolir dans l'ensemble, on ne peut guère faire usage de ces types généraux pour démontrer une « parenté de langues ». Car il arrive souvent que, avec le temps, le type tende à s'effacer plus ou moins complètement, ainsi qu'il ressort de l'histoire des langues indo-européennes.

L'indo-européen commun présentait de la manière la plus extrême le type qu'on nomme « flexionnel ». Tous les emplois des mots, tous les liens entre les mots se marquaient par des différences de la forme interne des mots. En sanskrit, où ce caractère de l' « indo-européen commun » est particulièrement bien conservé, « je suis » se dit *ásmi*, « ils sont » se dit *sánti*, « il a été » se dit *ā́sa*, et ainsi de suite. Si le mot « père » est sujet, il est de la forme *pitā́* ; s'il est complément direct, il est de la forme *pitáram* ; et, s'il est complément de nom, il est de la forme *pitúḥ*. Les différences sont profondes, on le voit, entre les formes d'un même mot ; et c'est au moyen de ces différences que s'indique tout le rôle des mots. Jusqu'à présent, les langues de la famille indo-européenne même les plus évoluées ont gardé quelque chose de ce type ancien ; en français, par exemple, le substantif est devenu invariable (l'-*s* du pluriel est purement graphique) ; mais le verbe comporte encore une forte part de flexion :

la différence entre *aime* et *aimez*, *j'aimais* et *nous aimions* est caractéristique, et plus encore celle entre *je veux*, *nous voulons*, *je voudrais*, *j'ai voulu*, etc. Toutefois la structure générale du français actuel diffère absolument de celle qu'avait l'indo-européen. Et celle de l'anglais en diffère plus encore. Les langues romanes, la plupart des langues germaniques, les langues iraniennes ne méritent plus vraiment aujourd'hui le nom de « flexionnelles ». Et même les langues indo-européennes les plus conservatrices ont un type tout autre que l'indo-européen commun. Les structures des langues indo-européennes actuellement parlées sont très différentes de la structure qu'avait l'indo-européen commun et de plus très différentes entre elles. Dès lors, ce n'est pas à la structure générale qu'on reconnaît une langue indo-européenne.

La langue « indo-européenne » commune procédait par suffixation, c'est-à-dire que les éléments grammaticaux affixés à la partie du mot exprimant le sens étaient toujours postposés. Ainsi en latin, où cet usage s'est bien maintenu, la marque *-ēs* du nominatif et de l'accusatif pluriel et la marque *-um* du génitif pluriel sont postposées à *patr- : pat-rēs* « les pères », *patr-um* « des pères ». En français, au contraire, il tend à se constituer des préfixes. Le latin distinguait *amō, amās, amat* par la fin du mot. Le français distingue *j'aime, tu aimes* (avec *-s* purement graphique), *il aime*, par les éléments *je, tu, il* que la grammaire traditionnelle qualifie de « pronoms », mais qui n'ont pas d'existence autonome et qui sont actuellement de purs signes grammaticaux préposés. A cet égard encore, le français procède autrement que ne faisait l'indo-européen.

Ce n'est donc pas avec de pareils traits généraux de structure, sujets à changer du tout au tout en l'espace de quelques siècles, et du reste comportant seulement des variations peu nombreuses, qu'on peut établir des parentés de langue. Qui comparerait le latin où l'on a *liber Petri* ou *Petri liber* (les deux ordres sont possibles), c'est-à-dire où le rapport entre « Pierre » et « livre » est marqué par la forme *Petri* du nom dont le cas sujet est *Petrus,* le cas complément direct *Petrum,* etc., avec des variations de la fin du mot, et le français, où le rapport entre le nom et son complément est indiqué par la particule préposée *de* et par la place de *de Pierre* ne trouverait entre les procédés employés par les deux langues aucun trait commun.

Ce qui est probant pour établir la continuité entre une « langue commune » et une langue ultérieure, ce sont les procédés particuliers d'expression de la morphologie. Par exemple, il n'est pas rare que le rapport de dépendance entre deux substantifs soit exprimé par une particule soit préposée, comme fr. *de,* soit postposée comme angl. *-s*. Mais le fait que cette particule est de la forme *de* ou de la forme *-s* est caractéristique ; car n'importe quel autre phonème pourrait aussi bien indiquer le rapport, si la tradition n'en avait décidé autrement. Dès lors le fait que le rapport est marqué par *de* préposé au complément est un caractère distinctif d'un parler français, et le fait que le rapport est marqué par *-s* postposé au complément un caractère distinctif d'un parler anglais. — Il faut ajouter que l'emploi de *de* peut disparaître d'un parler français ou celui de *-s* d'un parler anglais sans que, pour cela, ces parlers cessent d'être français ou anglais. Seuls, les faits positifs ont une valeur probante.

Les faits singuliers de cette sorte sont souvent stables. La prononciation peut se transformer, le vocabulaire peut changer, alors que ces traits demeurent. Ainsi, dans les parlers actuels de la France du Nord, les formes locales des mots se modifient pour être conformées aux formes françaises, le vocabulaire se renouvelle, et l'on tend en général à parler suivant l'usage français commun. Ce qui subsiste en dernier lieu, ce sont des particularités locales de la morphologie, ainsi la marque du masculin et du féminin par *i dit, a dit,* là où le français commun a *il dit, elle dit.* Les faits particuliers de cette sorte sont appris dès l'enfance, ils deviennent des habitudes dont on ne prend pas conscience, et ils sont susceptibles de demeurer alors que tout le reste se modifie.

Dès lors une langue à morphologie touffue et complexe, comprenant un grand nombre de faits particuliers, se prête bien à la démonstration des parentés, tandis qu'une langue à morphologie simple, opérant surtout avec des procédés généraux tels que l'ordre des mots, rend malaisée la découverte de preuves valables. On n'a presque pas besoin de démontrer qu'une langue est indo-européenne : partout où l'on a trouvé une langue indo-européenne encore inconnue, le « tokharien » ou le « hittite » dans les derniers temps, le caractère indo-européen s'en est révélé dès le début du déchiffrement et de l'interprétation. Au contraire, les langues d'Extrême-Orient qui, comme le chinois ou l'annamite, n'offrent presque pas de particularités morphologiques, n'ont par là même rien où puisse se prendre le linguiste qui essaie de trouver des langues parentes aux parlers chinois ou aux parlers annamites ; et la restitution d'une « langue commune » dont le chinois, le tibétain, etc., par exemple, seraient des formes postérieures, se heurte à des obstacles quasi invincibles.

Plus sont singuliers les faits dont on constate entre deux langues la concordance, et plus grande est la force probante de la concordance. Les formes anomales sont donc celles qui sont le plus propres à établir une « langue commune ».

Le fait que *il est, ils sont, je fus* du français concorde avec *est, sunt, fuī* du latin est de nature à faire apparaître que le français est une forme nouvelle prise par le latin. Des concordances comme *je veux, nous voulons, je voulus* avec *uolō, uolumus, uoluī,* ou comme *je dis, nous disons, dit* avec *dīcō, dīcimus, dictus* viennent confirmer la preuve. Les verbes « irréguliers » apportent ainsi à la démonstration des moyens de preuve en grand nombre.

L'adjectif confirme la démonstration. On y observe en effet la distinction d'un masculin et d'un féminin, dont l'emploi est arbitraire, quand elle ne dépend pas du sexe. Or, cette distinction dénuée de sens se trouve en français comme en latin : *veston neuf* et *veste neuve,* de même que le latin a *nouus, noua.* Il apparaît même dans certaines formes, l'article *la,* les adjectifs possessifs *ma, ta, sa,* que le féminin se caractérise parfois par *a* en français comme en latin.

Plus une langue comprend ainsi de catégories grammaticales non significatives ou de formes anomales, et plus la démonstration des parentés et la restitution d'une langue commune initiale sont faciles pour le linguiste.

La communauté initiale ne se reconnaît pas toujours à une conservation pure et simple ; elle se traduit souvent par des innovations divergentes. Ainsi l'indo-européen avait un certain type de flexion verbale de forme compliquée qui a tendu à s'éliminer partout ; on en reconnaît la présence ancienne à des innovations qui sont diverses suivant les langues. Par exemple il y avait un présent de la forme **eiti* « il va », **y-enti* ou **iy-enti*, **y-onti* ou **iy-onti* « ils vont » ; le sanskrit le garde sous la forme *éti*, *yánti* ; le grec attique l'a, un peu plus altéré, dans *eisi*, *iāsi* ; le latin, avec plus d'altération encore, dans *it*, *eunt* ; le vieux lituanien a le singulier *eiti* ; mais ces formes étaient trop singulières : le lituanien moderne a remplacé *eiti* par une formation de type nouveau, régulière en lituanien, *eina* ; le slave a, dès le IXe siècle après J.-C., dans la langue des anciens traducteurs, *idetŭ* « il va », *idǫtŭ* « ils vont ». Les formes nouvelles du lituanien, *einu* « je vais », et du slave, *idǫ*, réfléchissent à leur manière la forme ancienne qui s'est éliminée.

Dans l'exemple cité, la forme ancienne est conservée par certaines langues, et l'on s'explique aisément l'innovation de certaines autres. Mais il arrive que la forme ancienne ne soit conservée nulle part, et ce n'est que l'absence de concordance de détail des formes rapprochées qui laisse supposer une forme singulière éliminée par la suite. Il a dû y avoir en indo-européen un présent de la forme **melǝti* et **molǝti* « il moud » avec une 3e personne du pluriel **m°lenti* et **m°lonti*. Cette flexion difficile à manier s'est éliminée partout. On en reconnaît néanmoins l'existence ancienne à ce que le slave a *meljetŭ*, le lituanien *malu* et le gotique *maliþ*, l'irlandais *melid* et les langues brittoniques toutes voisines une forme à vocalisme différent *mal-*, etc. Nulle part, les formes indo-européennes supposées ici ne sont conservées ; mais on les restitue par la connaissance qu'on a du système indo-européen et par les divergences observées entre les langues historiquement connues.

L'observation des faits singuliers ne dispense pas d'examiner l'ensemble de la morphologie. Toute morphologie constitue un système complet. D'après l'ensemble des concordances, il faut donc restituer le système de la « langue commune », dans la mesure du possible. Et, ceci fait, il faut voir comment à ce système s'opposent les systèmes morphologiques, souvent tout différents, on l'a vu, des langues en lesquelles la langue initiale s'est transformée. Les particularités sont le moyen décisif de preuve. Mais la démonstration n'est définitivement acquise que lorsqu'on a confronté système morphologique à système morphologique et qu'on a vu comment il est possible de passer du système initial aux systèmes ultérieurs.

On observe souvent alors que des caractères généraux du système ancien se maintiennent longtemps. Ainsi, en indo-européen commun, où chaque mot portait en lui-même la marque de son rôle dans la phrase, le nom se distinguait nettement du verbe. Or, encore aujourd'hui, dans les langues indo-européennes, la distinction du nom et du verbe est demeurée nette partout.

Le phonétisme comporte aussi la mise en présence d'un système ancien et de systèmes nouveaux, tous ces systèmes pouvant différer les uns des autres aussi profondément qu'il est imaginable.

Mais les différences d'un système à l'autre ne sont pas capricieuses. S'il est souvent impossible de reconnaître des correspondances régulières entre deux des langues issues d'un original commun, les correspondances entre la langue commune et chacune de ces langues obéissent à des règles fixes et qui se laissent exactement formuler. C'est ce que l'on appelle « lois phonétiques ».

A *p, t, k* de l'indo-européen répondent en germanique *f, þ, x* (d'où *h*) et, en certaines conditions, *ƀ, đ, γ* ; à *b, d, g* de l'indo-européen, répondent germ. *p, t, k* ; à *bh, dh, gh* de l'indo-européen répondent germ. *b, d, g* (*ƀ, đ, γ* en position intervocalique). Ce système régulier de correspondances en germanique est ce que l'on appelle la « mutation consonantique », ou loi de Grimm (vue en grande partie par Rask, un peu avant Grimm, mais posée comme loi par Grimm). La régularité des correspondances entre la « langue commune » initiale et les langues ultérieures exprime le fait que les changements de la prononciation n'atteignent pas isolément tel ou tel mot, telle ou telle forme, mais qu'ils portent sur le système phonique même.

Du principe de la méthode, il résulte que des règles de correspondances peuvent être posées entre la langue initiale et chacune des langues qui la continuent, mais non entre les diverses langues continuant une même langue commune. On peut poser que *p* initial de l'indo-européen est représenté en grec et en sanskrit par *p*, en gotique par *f*, ainsi dans gr. *patḗr*, skr. *pitā́* et got. *fadar* « père », et que k^w initial de l'indo-européen devant *o* est représenté en grec par *p*, en sanskrit par *k* et en gotique par *hw*, ainsi gr. *póteros*, skr. *katarā́ḥ*, got. *hwaþar* « lequel des deux ». Les deux correspondances constatées en fait :

gr. *p* = skr. *p* = got. *f*
gr. *p* = skr. *k* = got. *hw*

ne sont intelligibles que par rapport à une forme indo-européenne commune supposée pour les expliquer.

Ce n'est pas avec des ressemblances de formes qu'on opère quand on compare des langues d'une même famille, mais uniquement avec des règles de correspondances. On a vu ci-dessus, p. 6, que le nom de nombre « deux » de l'arménien, qui est *erku,* répond à l'ancien **dwō* (ou **duwō*) de l'indo-européen. Cette correspondance semble au premier abord étrange. Mais il y a une règle générale en vertu de laquelle à *dw-* de l'indo-européen répond *erk-* de l'arménien. Car on connaît deux autres exemples de cette correspondance. L'indo-européen a eu une racine **dwei-* « craindre » qui est largement représentée en grec homérique par des formes verbales telles que **dedwoa* (noté *deidō*), *dedwoike* (noté *deidoike*) « je crains » ou nominales telles que **dweos* (noté *deos*) « crainte » ; l'arménien y répond par *erkiwł* « crainte », *erkeay* « j'ai craint ». Il y a eu un adjectif **dwāro-* « long », conservé notamment par le grec ; or, l'arménien a un adjectif *erkar* « long ». La règle est donc fondée sur trois rapprochements évidents. Si l'on pense au nombre restreint des mots indo-européens connus présentant l'initiale **dw-*, **duw-*; la coexistence de ces trois rapprochements apparaît probante.

La correspondance s'explique du reste. Le groupe de consonne dentale suivie de *w* aboutit en arménien à une gutturale : *tw-* est représenté par *k'-*, ainsi dans *k'o* « de toi », en face de formes telles que *twe* (d'où attique *se*) du grec. La sourde *k* provient de ce que, en arménien comme en germanique, il y a eu une mutation des anciennes occlusives sonores en sourdes : *d* passe à *t*, *g* à *k*. L'*r* qui précède est une trace de l'ancien caractère sonore de la consonne initiale du groupe (pour le détail de l'explication, voir M. Grammont, *Mémoires de la Soc. de ling.*, XX, p. 252), cette *r* a entraîné à son tour le développement de l'*e* initial qui figure dans *erku*. La présence de *-r-* à l'intérieur du mot a du reste suffi à empêcher le développement de *r* initiale : *kr-kin* est le mot qui signifie « double » (*krkin* repose sur un ancien **kirkin* ou **kurkin*). Tout, dans la correspondance si étrange de arm. *erk-* avec **dw-*, résulte de la structure de l'arménien. Une correspondance qui ne s'expliquerait pas ainsi serait suspecte.

Dans des groupes de langues où l'on a de longues séries de rapprochements clairs et dont par suite l'appartenance à une même famille ne fait pas question, on peut utiliser des correspondances surprenantes, telles que celle de arm. *erk-* en face de **dw-*, **duw-* des autres langues du groupe, et l'on arrive à pouvoir poser arm. *erk(i)-* = lat. *bi-*, si bizarre que cela paraisse au premier abord. Il va de soi que l'on ne peut se servir de pareilles correspondances pour commencer à établir une communauté de langues. On est obligé de commencer par des rapprochements dont l'explication apparaît aisément.

La régularité des correspondances n'exclut pas l'existence de traitements particuliers. Dans la phrase, les mots se trouvent en des positions diverses, en des conditions diverses. La régularité des traitements provient souvent de ce qu'il s'est fixé une forme moyenne entre des cas divers suivant les positions dans la phrase. Mais il y a notamment des cas où des formes prononcées plus rapidement ou avec peu de soin sont plus usuelles que d'autres, et il résulte de là des traitements spéciaux des mots accessoires : une réduction comme celle de *hiu tagu*

« ce jour-ci » à *hiutu* (all. mod. *heute*) en vieux haut allemand n'entre dans aucune catégorie générale.

Ces phénomènes qui ne se ramènent pas à des formules générales ne font pas difficulté dans une langue dont l'histoire est connue et pour laquelle on possède des règles de correspondances bien établies. Il est superflu de noter que, dans une langue dont la théorie n'est pas établie de manière exacte, des faits aussi aberrants sont inutilisables et qu'il convient d'en faire provisoirement abstraction.

Quant au vocabulaire, c'est dans la langue l'élément de tous le plus instable. Les mots sont sujets à disparaître pour les raisons les plus variées ; ils sont remplacés par de nouveaux termes. Au vocabulaire indigène il peut s'adjoindre des mots nouveaux aussi et plus nombreux que les anciens : c'est ainsi que, au vocabulaire germanique de l'anglais, s'est superposé un élément latin et français aussi considérable que l'élément germanique. Il arrive même que tout le vocabulaire appartienne à un groupe autre que celui dont relève la morphologie : le tsigane arménien a une morphologie et une phonétique tout arméniennes et un vocabulaire tout tsigane.

Là même où ils subsistent, les mots sont susceptibles de changer si fort de valeur que, malgré l'ancienneté de leur forme, ils sont vraiment des mots nouveaux. Le nom indo-européen du chef de famille indo-européen, le **pəter-*, a subsisté en français ; c'est le mot *père*. Mais au lieu que le mot indo-européen désignait un rôle social, et qu'on pouvait qualifier de **pəter-* le dieu suprême, comme le plus grand des chefs de famille, le mot français désigne de plus en plus « celui qui engendre », et il passe au sens de lat. *genitor*, presque au sens de « mâle » : on dit d'un lapin mâle que c'est un *père*. Ainsi fr. *père* continue un mot indo-européen, mais avec une valeur si différente de la valeur indo-européenne que c'est en réalité un mot nouveau.

Malgré cette instabilité fréquente du vocabulaire, ce sont les concordances de vocabulaire qui frappent d'abord quand on rapproche des langues entre elles. Souvent même on ne dispose que du vocabulaire, soit que les langues envisagées soient mal connues et qu'on n'ait de témoignages que sur le vocabulaire, soit qu'il s'agisse de langues à grammaire très simple, comme en Extrême-Orient, soit que les morphologies qui ont subsisté se soient constituées trop tard, après la période de communauté supposée. Il importe donc beaucoup d'examiner comment peut se prouver une concordance de vocabulaire.

Il a été noté déjà que les rapprochements étymologiques valables ne se font jamais d'après des ressemblances de forme phonétique, mais seulement d'après des règles de correspondances : si l'on peut rapprocher arm. *erku* de russe *dva*, ce n'est pas que les deux formes se ressemblent : les formes phonétiques n'ont rien de commun ; c'est que les règles de correspondances permettent le rapprochement, *ō* de l'indo-européen aboutissant à *a* en slave, à *u* en arménien, et *duw-* de l'indo-européen aboutissant à *dv-* en russe, *erk-* en arménien.

Pour autant qu'elles ne s'expliquent pas par des conditions propres à tel ou tel mot, les irrégularités dans les correspondances dénoncent soit un emprunt à d'autres langues soit une étymologie à écarter. Un ancien *ca-* du latin est représenté en français par *cha-*, *che-*, *chè-*, ainsi dans *campum* donnant *champ*, *carrum* donnant *char*, *caballum* donnant *cheval*, *carum* donnant *cher*, etc. ; si donc on a *camp* en regard de *campum*, c'est que le mot n'appartient pas à la tradition ancienne du français ; en fait il vient de l'italien, et l'on sait en effet quand et pourquoi il a été « emprunté ». C'est un *b-* qui en germanique peut répondre à *f-* du latin ; on a par exemple en latin *flōs*, *flōris* et en allemand *blume* ; dès lors *feuer* de l'allemand n'a aucun rapport avec *feu* du français : il suffit d'ailleurs de penser aux correspondants romans de fr. *feu*, à savoir it. *fuoco*, esp. *fuego*, pour que la ressemblance de *feu* et de *feuer* s'efface. On ne fait donc des rapprochements qu'avec des formules précises de correspondances — et en prenant soin d'éviter ce qui est dû à des emprunts.

Il est malheureusement difficile de déterminer a priori ce qui a chance d'être indigène et ce qui a chance d'être emprunté. Les verbes et les adjectifs sont plus ordinairement indigènes que les substantifs : des verbes comme *vivre* et *mourir*, *venir* et *dormir* du français actuel sont encore des verbes indo-européens, et de même des adjectifs comme *vif*, *vieux*, *neuf*, etc. Néanmoins un verbe fort comme all. *schreiben* est emprunté à lat. *scrībere*. Inversement beaucoup de substantifs français tels que *pied*, *chien*, *veuve*, etc. sont indo-européens. — Le risque qu'un mot soit emprunté est toujours grand, et l'étymologiste, d'une langue ancienne ou récente, qui raisonne comme si les mots à expliquer avaient a priori toutes chances d'être indigènes s'expose à des erreurs fréquentes.

Là où il s'agit de mots remontant vraiment à la « langue commune », il faut restituer un mot de cette langue défini à tous égards, et ne pas se contenter de rapprocher de petits éléments radicaux. Et, comme les risques d'erreur sont grands, il est nécessaire de s'assurer, par des précisions, que les concordances observées ne sont pas fortuites.

Le premier point, sur lequel on est d'accord en fait sinon en principe, c'est qu'une étymologie est valable seulement si les règles de correspondance phonétique sont appliquées d'une manière exacte, ou, au cas où une divergence est admise, si cette divergence est expliquée par des circonstances particulières rigoureusement définies.

Il va de soi que la concordance risque d'autant moins d'être fortuite que le nombre des éléments phoniques qui se correspondent est plus grand.

Le hasard ne peut évidemment pas faire que « veuve » se dise *vidhávā* en sanskrit, *vĭdova* en slave, *widdewū* en vieux prussien, *widuwo* en gotique, *fedb* en irlandais, *uidua* en latin, étant donné que, par exemple, *f* de l'irlandais et *u-* (consonne) du latin répondent au *v* initial du sanskrit. La concordance de *w*, de *i*, de *dh* et de *u*, dans cet ordre, ne saurait être accidentelle. Entre les formes sanskrite, slave, prussienne, d'une part, et les formes germanique, irlandaise et sans doute latine, de l'autre, il y a, il est vrai, une différence : la présence d'une voyelle *e* (ou *o*) entre *dh* et *w* dans la première langue et l'absence de cette voyelle dans les autres ; cette différence ne fait pas difficulté ; car elle relève des alternances vocaliques connues de l'indo-européen.

Mais si, au lieu de comparer ainsi quatre éléments phoniques se correspondant exactement, on n'en compare que trois, la preuve est moins forte, pour devenir fragile si l'on dispose seulement de deux éléments concordants, et à peu près nulle si l'on dispose d'un seul.

Aussi les linguistes qui opèrent avec de petits éléments radicaux et qui même souvent analysent les racines pour ne plus rapprocher que des fragments de racines ruinent d'avance par là la démonstration qu'ils cherchent. Un rapprochement de mots complets ayant quelque étendue peut être sûr. Un rapprochement qui porte seulement sur une ou même deux consonnes radicales est sans valeur s'il n'est pas appuyé par des faits tout particuliers.

La concordance de sens doit être aussi exacte, aussi précise que la concordance de forme phonique (suivant les règles de correspondance). Ceci ne veut pas dire que les sens doivent coïncider plus que le phonétisme ; seulement les divergences des sens, s'il y en a, doivent s'expliquer, non par des possibilités vagues et générales, mais par des circonstances particulières. Le fait que le fr. *ouaille* repose sur lat. *ouicula* n'est pas rendu douteux par le fait que, dans le français actuel, ce mot désigne seulement le fidèle du prêtre de telle ou telle église. Mais le rapprochement se justifie parce qu'on sait que les fidèles d'une église chrétienne sont couramment comparés à un troupeau que fait paître le pasteur de cette église. Dans les parlers locaux français, le mot *ouaille* se trouve du reste au sens de « brebis », ce qui achève de démontrer la correction de l'étymologie.

En matière d'étymologie indo-européenne, le fait que les mots ont des flexions variées fournit des confirmations décisives. Soit, par exemple, le mot qui signifie « mouton » (mâle ou femelle : l'indo-européen n'a pas de noms distincts pour les animaux suivant le sexe) : sanskrit *áviḥ*, grec *óis* (ainsi chez Homère), lat. *ouis*, lituanien *avis* ; la concordance de sens et de forme entre toutes ces langues est parfaite ; et, comme on arrive à restituer ainsi un mot bien défini par la forme phonique et par le sens, l'étymologie peut passer pour assurée déjà. Mais il s'ajoute encore ceci que les deux langues les plus anciennement attestées, celles qui par là sont propres à donner l'idée la plus juste de certains détails des faits indo-européens, ont une flexion de type particulier : le génitif-ablatif sanskrit *ávyaḥ* concorde exactement avec grec *oiós*, qui repose sur **owyós*. Dans les autres langues, la flexion a été normalisée, ainsi qu'on l'attend. Mais le sanskrit et le grec fléchissent le mot de même. C'est une précision supplémentaire qui achève d'exclure le hasard et ajoute à la rigueur de la démonstration.

Quand on doit restituer une « langue commune » initiale, il y a lieu de tenir compte du nombre de témoignages qu'on a pour un mot donné. Une concordance de deux langues, si elle n'est pas totale, risque d'être fortuite. Mais, si la concordance s'étend à trois, quatre ou cinq langues bien distinctes, un hasard devient moins vraisemblable. Bien que *rādiy* « à cause de » du vieux perse et *radi* « à cause de » du slave ne se retrouvent pas ailleurs, on n'hésite pas à rapprocher les deux mots parce que forme, sens et détail de l'emploi concordent de tout point. A part un cas de ce genre, un rapprochement de mots trouvés dans deux langues indo-européennes seulement est suspect, à moins qu'on ne puisse reconnaître des conditions spéciales ayant déterminé l'élimination du mot dans les autres langues.

De quelque langue qu'il s'agisse, une étymologie ne peut passer pour prouvée que si un ensemble de concordances précises établit que les ressemblances des mots rapprochés ne peuvent être dues au hasard.

Tout mot n'a pas droit à une étymologie, comme on le croirait parfois à parcourir les dictionnaires étymologiques. La règle de méthode est que seuls des faits positifs bien déterminés justifient un rapprochement : la preuve que le hasard est exclu est à la charge de l'étymologiste. Cette preuve peut être administrée de façons diverses suivant les langues et suivant les cas; les conditions en varient d'une langue à l'autre, d'un mot à l'autre. Mais toujours il faut que les preuves soient précises, rigoureuses.

Aussi le nombre des bonnes étymologies est-il assez restreint, au moins pour les langues dont l'histoire n'est connue que par comparaison et où la méthode comparative ne trouve pas dans des séries de textes d'époques diverses l'appui de témoignages historiques précis. Quand on lit un manuel de grammaire comparée, on s'étonne d'abord de le trouver bâti sur un petit nombre de rapprochements. C'est qu'il y en a peu sur lesquels il soit permis de faire fond sans réserve.

Une langue comme l'indo-européen, où les mots sont le plus souvent au moins dissyllabiques ou trisyllabiques et où les flexions sont variées et complexes, se prête bien à des démonstrations étymologiques parfaites (on peut négliger les fantaisies, plus ou moins laborieuses, des linguistes qui se contentent à bon marché). De même le sémitique, avec ses racines qui comprennent normalement trois consonnes, permet des étymologies parfaites. Au contraire, les langues où les mots sont courts, souvent monosyllabiques, et ne comportent pas de flexions particulières, excluent, par leur structure, les démonstrations étymologiques rigoureuses. Il y aura là une méthode nouvelle à trouver si l'on veut parvenir à de véritables démonstrations.

On n'a envisagé ici que les étymologies faites par comparaison entre langues diverses et reportant à un mot de la langue commune initiale. Pour déterminer de quelle manière peut se faire l'histoire de tous les mots — ce qui est hors du cadre de ces conférences —, il faudrait une longue étude, et qui envisagerait des procédés de toutes sortes.

Par exemple, il y a des mots qui s'expliquent, à l'intérieur d'une langue donnée, par des procédés de formation propres à cette langue. Un nom comme *(w)érgon* « œuvre » du grec trouve une étymologie par le rapprochement avec *werk* de l'allemand et par l'observation que la forme de ce mot est normale en indo-européen, la racine se retrouvant du reste en arménien et en iranien. Mais, pour expliquer *órganon* « instrument » du grec, il faut envisager un suffixe courant qui a fourni au grec bien d'autres noms.

Il y a beaucoup de mots dont l'histoire ne s'explique que par des données de fait : ainsi *organum* du latin n'est explicable que si l'on sait que le vocabulaire technique du latin est en grande partie grec ou imité du grec. Et, si l'on a en français deux représentants de ce mot, *orgue* et *organe,* avec des sens qui n'ont rien de commun ; ceci résulte de conditions spéciales ; si ces conditions sont connues, c'est parce qu'il s'agit de faits ayant eu lieu durant des périodes historiques ; aucune méthode comparative ne permettrait de les découvrir.

En fait, la plus grande part de l'étymologie des mots échappe à la méthode comparative et relève de l'histoire faite à l'aide de témoignages. Là où les témoignages manquent, et où l'on ne peut suivre, en se servant de données positives et précises, les accidents survenus à un mot, il est le plus souvent vain d'essayer d'en donner une véritable étymologie.

Qu'il s'agisse de morphologie, de phonétique ou de vocabulaire, le principe qu'il ne faut jamais perdre de vue, c'est que les rapprochements valent seulement dans la mesure où ils sont soumis à des règles strictes. Plus le linguiste se donne de liberté, plus ses rapprochements comportent d'arbitraire, et plus ses démonstrations sont précaires. Par exemple, le linguiste qui interprète des noms propres d'un pays par des rapprochements avec un idiome supposé parent de celui qui avant un changement de langue se parlait dans ce pays joue un jeu dangereux. Si l'on est sûr de la langue qui se parlait anciennement et si l'on connaît bien des langues toutes proches de celle-ci, quelques explications de ce genre sont sûres : une localité gallo-romaine nommée *Brīua* (*Brive*) a certainement un nom gaulois parce qu'elle est en un point où un pont a dû toujours exister, que les langues celtiques conservées supposent un nom *brīwā* du « pont » et que le latin a pris la place du gaulois, langue sûrement celtique, dans le pays où est Brive. Mais il est risqué d'expliquer par un rapprochement avec des langues caucasiques tel nom de localité grecque, dans un cas où rien n'avertit des sons que doit avoir étymologiquement le nom propre envisagé, où les langues rapprochées ont une forme non identique et où aucune donnée historique ne prouve qu'il ait été parlé une langue prochement apparentée aux langues caucasiques. D'une manière générale, les étymologies de noms propres sont incertaines parce que, des deux données dont la concordance avec des faits d'autres langues établit la valeur, le sens et la forme phonique, on peut utiliser un seul : la forme phonique. Les linguistes qui s'intéressent surtout à l'étymologie des noms propres sont souvent des aventuriers de la linguistique, et peu nombreux sont ceux qui ont toutes les exigences de méthode qu'il faut.

IV.

DU DÉVELOPPEMENT LINGUISTIQUE ENTRE L'ÉPOQUE DE COMMUNAUTÉ ET L'ÉPOQUE HISTORIQUE.

La comparaison permet, dans la mesure indiquée, de « restituer » la « langue commune » dont les langues d'une même « famille » sont les formes prises au cours du temps. Mais il reste à se rendre compte de ce qui s'est passé entre l'époque de « communauté » et la date où les langues commencent d'être attestées en fait. Le problème est difficile, souvent insoluble.

Il va sans dire qu'on doit s'aider des périodes intermédiaires de communauté qu'il est possible soit de déterminer exactement soit d'apercevoir. On restreint ainsi, autant qu'on le peut, l'importance de l'intervalle entre la « langue commune » et les langues attestées. Mais, ainsi qu'on l'a vu, p. 16 et suiv., ces périodes de communauté n'existent pas toujours, ou, quand elles ont existé, il n'est pas toujours possible d'en tirer parti ou d'en tirer parti pour tous les faits. Même dans les meilleurs cas, l'hiatus entre la « langue commune » et les langues attestées reste grand, et les interpolations qu'on fait pour le combler comportent une large part d'inexactitude et d'arbitraire.

La considération des faits linguistiques ne suffit pas à fournir même un commencement de chronologie. Car la rapidité des changements linguistiques varie dans une mesure étendue. Certaines langues peuvent demeurer de longs siècles presque sans changer : connus depuis le XII^e^ siècle, les parlers turcs ont gardé depuis lors les mêmes traits essentiels, si bien que, quoique séparés les uns des autres depuis un millier d'années parfois, ils continuent d'être très pareils les uns aux autres. De l'histoire des parlers polynésiens, on sait peu de chose; mais ce qui est frappant, c'est que, malgré les grandes distances qui séparent les unes des autres les îles de la Polynésie, ces parlers sont demeurés presque identiques les uns aux autres. Les cas de ce genre ne sont pas rares : les langues de peuples peu civilisés ont souvent une grande stabilité.

La rapidité de renouvellement qu'on prête parfois aux langues de demi-civilisés s'explique, en partie au moins, par des usages relatifs au vocabulaire. Il arrive souvent que, chez les demi-civilisés, des mots soient frappés d'interdiction ; par exemple à la suite de la mort d'un individu, le mot qui figure dans son nom est souvent interdit. Dès lors, le voyageur qui a connu un vocabulaire donné, peut, s'il revient quelques années après, en trouver un autre différent sur bien des points. M. Thalbitzer a fait connaître, récemment, un cas de ce genre chez des Eskimo. Mais il ne résulte pas de là que le fonds de la langue change pour autant.

Ailleurs, au contraire, on trouve des changements rapides du système linguistique. Au III^e^, et même au V^e^ siècle ap. J.-C., le latin avait encore sans doute beaucoup de son caractère ancien, surtout à première vue. Au IX^e^ siècle, les langues romanes présentaient déjà leurs traits essentiels ; et même, dès le vi^e siècle, la façon dont les écrivains de l'époque mérovingienne écrivent le latin indique souvent que les principaux changements étaient réalisés ou près de l'être. De même le changement assez profond qui, de l'arménien ancien, a fait l'arménien moderne s'est réalisé entre le v^e siècle au plus tôt et le x^e au plus tard ; depuis lors il n'y a plus eu en arménien de changement aussi profond. Au vi^e siècle av. J.-C., le perse avait encore un aspect archaïque ; au i^{er} siècle de notre ère, il avait déjà un aspect moderne, et il n'a subi depuis lors aucun changement dont l'importance soit comparable.

Le temps n'est donc que l'une des conditions d'où dépend l'importance des changements. Ni un bon état de conservation de l'usage ancien ne prouve que le temps écoulé depuis la rupture de la communauté initiale soit bref, ni un renouvellement étendu des formes ne suppose un long intervalle de temps. Et, comme les faits linguistiques ne fournissent pas d'autre indice, on ne saurait le plus souvent rien dire, même approximativement, du temps qu'ont pu demander les innovations.

Le degré de nouveauté d'une langue par rapport à la « langue commune » n'est pas reconnaissable par la date : parmi les langues sémitiques, il y en a deux qui ont conservé la déclinaison. Or, l'une de ces deux est l'akkadien du début du second millénaire avant le Christ, l'autre est l'arabe du vii^e siècle ap. J.-C. Dans des langues attestées bien avant l'arabe, comme l'hébreu ou l'araméen, il n'y a plus la déclinaison des noms.

S'il y a des moments où le changement linguistique semble se précipiter, cela tient pour beaucoup à ce qu'il y a, pour les innovations, une longue période de préparation. La réalisation du changement ne fait dans bien des cas que manifester l'aboutissement d'un long travail.

Un des principaux faits auxquels est dû le passage du type indo-européen ancien au type moderne des langues de la famille est la perte de l'accent de hauteur et du rythme quantitatif. Dans le type ancien, dont les textes védiques et grecs classiques donnent encore une idée nette, le « ton » consistait en une élévation de la voix et ne servait pas de sommet rythmique au mot ; il jouait un rôle sémantique analogue à celui que jouent les tons, c'est-à-dire les montées et descentes de la voix, dans les parlers chinois ou soudanais. Le rythme reposait tout entier sur l'alternance de syllabes longues et de syllabes brèves. La comparaison des diverses langues montre que ce type a été indo-européen commun. Ce type phonique s'est éliminé partout, et notamment en grec et en latin à l'époque impériale. Le rythme a cessé d'être fondé sur l'alternance des syllabes brèves et des syllabes longues ; une partie des longues se sont abrégées, ou même toutes les voyelles ont perdu les oppositions de quantité brève et longue. Et l'ancien ton est devenu le centre du mot, le sommet rythmique. Dès lors, le mot est « centré » autrement qu'il ne l'était, et toute la prononciation change de caractère. Quand cette transformation s'achève, l'ancienne syllabe tonique acquiert une importance particulière, tandis que les autres éléments du mot deviennent relativement secondaires. Dans les langues romanes, on sait que le traitement de la syllabe accentuée diffère fondamentalement de celui des autres syllabes du mot. Le changement, à peu près réalisé dès le IX^e^ siècle, de lat. *caballum* en fr. *cheval* est spécifiquement français ; mais il n'est qu'une conséquence de l'élimination du rythme quantitatif et du changement de caractère de l'ancien ton qui ont eu lieu durant la période impériale du latin, et qui, en vérité, avaient commencé, en quelque mesure, dès avant les premiers monuments écrits du latin, c'est-à-dire, dès avant le IIIe siècle av. J.-C., comme l'indiquent diverses circonstances. Le changement de *caballum* en *cheval,* qui s'est fait assez vite, est ce qui se voit. La préparation de ce changement a demandé de longs siècles et ne se voit guère.

Ces périodes de préparation ont existé sur tout le domaine indo-européen, presque partout avant l'époque historique, ou hors des moments où le développement est observable. L'histoire du grec et celle du latin donnent une idée de ce qui a pu se passer ; il ne faut pas juger de tout par là ; car, sur le domaine indo-iranien, par exemple, le développement du rythme semble avoir été assez différent. Mais, en gros, il y a eu partout des préparations de ce genre.

Ce que l'on observe en général dans les langues indo-européennes, c'est le résultat de débâcles brusques, consécutives à des périodes de préparation, débâcles à la suite desquelles la langue offre un aspect nouveau, différent du type indo-européen.

La flexion nominale tend à s'éliminer sur tout le domaine indo-européen. Cette élimination a été favorisée par la tendance de la fin de mot à se réduire, d'où il résultait que les caractéristiques des formes casuelles, situées dans la fin de mot, perdaient progressivement de leur clarté. Mais l'essentiel, c'est que la flexion casuelle s'est trouvée peu à peu remplacée par d'autres procédés d'expression. Dès avant la période historique, les formes casuelles ne suffisent plus à indiquer les relations locales : si le latin a obligatoirement une préposition dans *eo in urbem, habito in urbe, uenio ex urbe,* ce n'est pas chose isolée ; des faits analogues s'observent dans la plupart des langues indo-européennes dès le début de l'époque historique, et la pleine valeur des cas locaux ne s'est maintenue que dans des survivances plus ou moins isolées. Cet état une fois atteint, la flexion casuelle devenait une complication superflue, et il n'est pas surprenant que la langue l'ait ou restreinte ou supprimée partout.

Ce qui rend malaisé, souvent impossible, de suivre le développement de la langue entre l'époque de communauté et les époques historiques, ce n'est pas seulement la difficulté d'évaluer le temps qu'a demandé le changement, ni le caractère invisible, ou presque, des périodes de préparation des changements.

Il y a bien d'autres causes d'embarras.

Les changements ultérieurs sont largement commandés par l'état de la langue commune et par la façon dont elle se brise. Il suit de là que des changements identiques ou semblables ont lieu même après la séparation et le commencement de la différenciation des langues issues de la « langue commune ». Ce fait est souvent méconnu. Les traités de grammaire comparée procèdent souvent comme si tous les faits superposables des divers représentants d'une même « langue commune » remontaient à l'époque d'unité. Sans doute, les auteurs ne le croient pas au fond, et ils se garderaient assurément de l'affirmer ; ils indiquent même parfois des réserves à cet égard. Mais l'exposé est présenté comme si les auteurs admettaient pareille hypothèse. Tel est le cas notamment pour le remarquable *Grundriss* de Brugmann. Or, rien n'est moins conforme à la réalité.

Soit par exemple la caractéristique *-m* de la 1^re^ personne du singulier des verbes dans les langues slaves ; cette caractéristique se trouve dans une notable partie de la conjugaison dans la plupart des langues slaves modernes, et même dans tous les présents normaux en serbo-croate. Or, en slave commun, elle était propre à quatre ou cinq verbes anomaux. Ce n'est qu'au cours du moyen âge, à date historique, que l'innovation a eu lieu, en tchèque et en serbo-croate notamment. A cette date, les langues slaves étaient séparées depuis longtemps, et c'est de manière indépendante que chacune a réalisé l'innovation, dont, au surplus, les limites varient d'une langue à l'autre.

Des faits pareils s'observent ailleurs pour la même 1^re^ personne du singulier : l'arménien a aussi généralisé une forme caractérisée par *-m*. Ici, on n'a que l'aboutissement du procès, et le détail des faits par lesquels l'arménien y est parvenu échappe. — Dans l'Inde et dans l'Iran, il y a eu la même innovation, et il se faut de peu qu'on ignore tout de la manière dont elle s'est produite. Par bonheur, on possède une petite série de vieux textes iraniens, les *gāthā* de l'Avesta, où la généralisation de la désinence *-mi* n'est pas faite ; on voit par là que quoique le sanskrit ait pour dire « je porte », une forme *bhárāmi* et l'Avesta récent une forme *barāmi*, l'addition de *-mi* est postérieure à la période de communauté indo-iranienne (aryenne) : car les *gāthā* de l'Avesta ont encore le type *barā*.

Les concordances entre les formes historiquement attestées proviennent, beaucoup plus qu'on ne l'imagine au premier abord, de développements parallèles entre des langues déjà séparées et différenciées.

Si des créations semblables se réalisent ainsi d'une manière parallèle, mais indépendamment, certains traits anciens se perdent parallèlement, à bien plus forte raison. Grâce au fait que le grec connaît encore à l'époque classique la fameuse règle *ta zōa trekhei* (coexistence d'un verbe au singulier avec un sujet au pluriel neutre) et que la même règle se retrouve, exactement observée, dans les *gāthā* de l'Avesta, on sait que la forme dite « nominatif-accusatif pluriel neutre » de l'indo-européen est, en réalité, un ancien collectif. Mais une langue aussi archaïque que le védique l'ignore déjà presque totalement. Le latin a généralisé l'emploi du verbe au pluriel avec sujet au pluriel neutre ; le slave, le celtique et le germanique de même. Et, même en grec, la particularité s'est perdue avec le temps. Là où l'on n'a pas de formes très archaïques des langues — et c'est le cas ordinaire —, une large part des événements de la période intermédiaire demeure donc inconnue.

Mais il n'y a pas seulement des éliminations d'usages anciens et des innovations d'usages destinés à durer. Le développement linguistique est chose complexe, et l'on ne saurait avoir l'illusion ni d'expliquer tout ce qui s'est créé durant les périodes intermédiaires ni de restituer les événements multiples qui ont eu lieu durant ces périodes.

On suit, par des textes, l'histoire du grec depuis le VII^e siècle av. J.-C. jusqu'à l'époque actuelle, et l'on a là un exemple qui permet de juger la suite des créations et des pertes par lesquelles la langue est parvenue à l'état moderne.

Ainsi, avant l'époque historique, le grec a donné à chaque verbe un futur; puis, à l'époque byzantine, il a éliminé ce futur, et l'a remplacé par quelque chose de tout nouveau.

Le grec avait hérité d'un type de parfaits : avant et pendant l'époque historique, il en a élargi l'emploi, multiplié les formes ; il a créé des procédés grâce auxquels presque chaque verbe a pu recevoir un parfait. A l'époque attique, le parfait était une forme essentielle et couramment employée du verbe. Puis l'usage en a disparu, et le grec moderne n'a plus aucune forme personnelle du parfait.

Si donc on ne connaissait le grec, comme on connaît le slave, qu'à partir du ɪxᵉ siècle ap. J.-C., le linguiste ne soupçonnerait même pas que, avant d'arriver à l'état actuel du verbe à deux thèmes, le grec a traversé une époque où un futur et un parfait ont eu un riche développement, qu'il y a eu des créations multiples, des formes faites et refaites à plusieurs reprises : de tout ce développement, le linguiste qui ne verrait que l'aboutissement final n'aurait aucune idée.

Or, la plupart des langues indo-européennes sont connues seulement après le début de l'ère chrétienne, en un temps où sont achevés ces développements complexes que des données positives permettent d'observer en grec, et où, par suite, le linguiste ignore ce qui s'est passé réellement.

Dès lors, il serait vain de prétendre expliquer à coup sûr toutes les formes nouvelles qui se sont constituées entre l'époque de communauté et les époques historiques. On ne doit pas être surpris que la linguistique puisse faire seulement des hypothèses incertaines et fragiles sur l'origine du perfectum latin en *-uī*, du parfait passif grec en *-thēn*, du prétérit gotique en *-da* (*-ta*), etc. Ces formes résultent d'un travail qui s'est produit durant des périodes dont l'état de langue n'est pas restituable ; les données manquent pour déterminer le détail des conditions grâce auxquelles les innovations se sont réalisées. Vouloir les expliquer à tout prix, imaginer qu'on peut trouver ici des démonstrations rigoureuses, c'est perdre de vue les limites de la méthode comparative.

Entre l'époque de communauté et les époques historiquement attestées, il y a des développements non entièrement restituables et où il demeure nécessairement d'autant plus d'inconnu que la période intermédiaire est plus longue et surtout plus chargée d'innovations.

V.

LES DIALECTES.

Jusqu'ici on a envisagé la langue initiale et les langues historiquement attestées comme si c'étaient autant d'unités. Tel n'est pas le cas. La notion de « dialecte », qui intervient sans cesse en linguistique, en fait foi.

Il faut ici faire abstraction du sens vulgaire, qui oppose le « dialecte » à une langue littéraire établie.

Le nom est grec, et il a été fait pour des situations grecques. Le grec ancien n'est pas, comme le latin, une unité. Il a été écrit de manières diverses suivant les temps et suivant les lieux. Dans tel type de parlers, on dit *nāsos* « île » et dans tel autre, *nēsos* ; dans tel, le datif pluriel est *possi, posi* « aux pieds » et, dans tel autre, *podessi* ; dans tel, « je veux » se dit *boulomai,* dans tel autre *bellomai,* et, dans tel autre, *dēlomai* ; dans tel, « possession » se dit *ktēsis,* et, dans tel autre, *p(p)āsis*. Mais ces différences ne sont pas telles que des Grecs employant des parlers de types différents aient eu le sentiment de parler des langues distinctes. Certaines de ces différences, comme l'opposition de *nēsos* à *nāsos* entrent dans de grandes séries de correspondances phonétiques, d'autres comme celle de *podessi* à *possi, posi* relèvent de différences grammaticales régulières, d'autres enfin sont de simples faits de vocabulaire, ainsi *ktēsis* : *(p)pāsis*. Ces diversités n'empêchaient pas les Grecs de sentir qu'ils avaient tous une même langue dans l'ensemble. Ce qui caractérise d'abord le « dialecte », c'est donc la diversité dans l'unité, l'unité dans la diversité.

Le « dialecte » même n'est pas un. Les Grecs distinguaient trois grands dialectes : ionien, éolien et dorien, auxquels les linguistes actuels ajoutent un quatrième, l'achéen (arcado-cypriote). Mais un texte donné n'est pas écrit en un dialecte ; il y a seulement des formes diverses de chaque dialecte. Par exemple, il n'y a pas un parler dorien : il y a le corinthien (avec le syracusain), l'argien, le laconien, le crétois, le rhodien, etc. Et encore ces types généraux admettent des variétés : il n'y a pas un parler crétois, et l'on observe des différences appréciables d'une localité de la Crète à une autre. Un « dialecte » est un ensemble de parlers divers présentant des caractères communs et plus semblables entre eux qu'ils ne sont aux autres parlers de la même langue.

Troisième point : les dialectes n'embrassent pas nécessairement tous les parlers d'une langue. Ainsi, en grec, le parler de tous le plus connu, l'attique, offre avec l'ionien des particularités communes, mais il ne fait pas partie proprement du dialecte ionien, et, moins encore, d'aucun autre. Les parlers de l'Elide, de la Locride, de la Phocide sont, à bien des égards, proches du dorien ; mais ils ne sont pas proprement doriens. Les parlers d'une langue ne se laissent donc pas toujours grouper tous en dialectes définis.

Ainsi la notion de dialecte est fuyante. Elle l'est d'autant plus qu'elle résulte de conditions historiques diverses.

Les dialectes grecs proviennent de ce qu'il y a eu, sur le domaine hellénique, des poussées successives et différentes d'envahisseurs. La plus ancienne vague que l'on observe est peut-être celle des Achéens, dont on n'a pas l'histoire, mais que les documents hittites récemment découverts laissent entrevoir vers le XIII^e siècle av. J. C. et dont le souvenir survit, sous forme légendaire, dans les poèmes homériques ; il en reste à date historique les parlers arcadiens dans le Péloponnèse, le pamphylien au Sud de l'Asie-Mineure, et le cypriote, qui jusqu'au bout a marqué un terme extrême de l'expansion des Hellènes. La vague ionienne est sans doute à peu près de même époque ; il n'en reste rien dans la Grèce propre ; mais l'Eubée, les îles de la mer Egée, et les rives de la Méditerranée, de la mer Noire aux côtes de Gaule, ont des colonies ioniennes. La plus récente de ces grandes poussées est celle des Doriens qu'on aperçoit encore en progrès aux débuts de la période historique de la Grèce et qui a recouvert nombre de régions déjà entièrement occupées par des hommes parlant d'autres dialectes helléniques, notamment une grande partie du Péloponnèse et la Crète.

Les dialectes qu'on rencontre sur le domaine galloroman ont une toute autre origine. Tous proviennent de la transformation d'un idiome sensiblement un employé en Gaule à la suite de la conquête romaine. Les parlers qui sont résultés de cette transformation ont des aspects divers en Provence et en Gascogne, en Normandie et dans l'Ile-de-France. On ne sait bien ni à quoi est due cette diversité, ni d'où vient que ces régions ont des parlers semblables entre eux quoique distincts. Comme les limites de ces régions linguistiques concordent souvent avec d'anciennes limites administratives, on a parfois supposé que les divisions politiques du pays conditionnent les divisions linguistiques. Mais si le pays a été ainsi divisé, c'est en conséquence des relations établies entre les hommes ; et si les divisions administratives des Romains ont été maintenues en gros par l'église chrétienne, c'est qu'elles répondaient en effet à des besoins naturels. Ce sont sans doute ces relations qui rendent compte des ressemblances de parler. Il y a coïncidence, et non cause.

Il faut tenir compte d'un autre type de faits. Dans le monde moderne, et depuis longtemps, il y a des centres d'où la civilisation se répand sur toute une région. En France, les centres provinciaux n'ont qu'une importance secondaire, et l'influence de Paris a dominé de bonne heure. En Italie, au contraire, il y a plusieurs centres provinciaux puissants. L'importance économique et politique de Venise est pour beaucoup dans la constitution d'un dialecte vénitien.

L'existence de dialectes a des conséquences importantes pour le comparatiste ; mais ces conséquences diffèrent suivant la manière dont les dialectes se sont établis.

Parfois la langue commune était une autant qu'une langue peut l'être. C'est, pour les langues romanes, le cas du « roman commun » qui est le parler courant d'une ville unique, Rome. Tout au plus y a-t-il à tenir compte, suivant l'hypothèse de Gröber, de la date où le latin a été porté à la région où il s'est étendu : le sarde aurait gardé la distinction de *ĭ* bref et de *ē* long, de *u* bref et de *ō* long, parce que, à la date où la Sardaigne a été conquise, Rome distinguait encore *ĭ* et *ē*, *ŭ* et *ō* ; le gallo-roman (ainsi que la plupart des parlers romans) aurait un représentant unique de *ĭ* et de *ē*, de *ŭ* et de *ō*, parce que, à la date où a eu lieu la conquête et où le latin s'est établi, la tendance à confondre *ĭ* et *ē*, *ŭ* et *ō* existait déjà. Mais, en somme, l'original des langues romanes est une langue une.

Il en va autrement de l'indo-européen commun. Plusieurs des différences qu'on observe entre les langues indo-européennes se trouvent à la fois dans des langues voisines des unes des autres, de telle sorte qu'on est conduit à poser que le point de départ de ces différences se trouve dans des différences existant déjà en indo-européen.

Soit par exemple le nom de nombre « dix » en grec, *déka*, et en latin, *decem*, d'une part — en sanskrit, *dáça*, et en arménien, *tasn*, de l'autre. Les formes germaniques et celtiques reposent sur des formes comportant un *k* médian ; les formes iraniennes, slaves et baltiques, sur des formes comportant une sorte de sifflante intérieure. A ce point de vue, il y a donc deux groupes de parlers indo-européens. Or, les mêmes parlers ont pour l'interrogatif-indéfini, dans les groupes grec, latin, celtique, germanique, un *qu* ou les représentants d'un *qu*, soit latin *quis*, et, dans les groupes indo-iranien, slave et baltique, arménien, un *k* ou le représentant d'un *k*, ainsi en sanskrit *káḥ* « qui », en lituanien *kas*, etc. Ceci tendrait à faire supposer qu'il y aurait eu, en indo-européen, deux groupes dialectaux, l'un représenté par le grec, l'italo-celtique et le germanique, l'autre par le slave, le baltique, l'arménien et l'indo-iranien. Au point de vue des consonnes dites gutturales, cette répartition serait exacte, et l'on parle souvent, pour l'indo-européen, d'un groupe de *satəm* (d'après une forme iranienne du nom de nombre « cent ») et d'un groupe de *centum* (d'après la forme latine).

Mais, si l'on examine d'autres faits, on trouve d'autres répartitions. Ainsi la distinction de *o* et de *a* est conservée en arménien, en grec, en italo-celtique ; mais elle est effacée en germanique, en baltique et en slave, en indo-iranien. Cette fois, le germanique marche avec le slave, etc., et l'arménien avec le grec, etc.

On n'arrive donc pas à poser, à l'intérieur de l'indo-européen, des dialectes définis, mais des différences de traitement pour certains faits phonétiques, morphologiques ou lexicaux. Et chacune de ces différences a ses limites propres. Le domaine de l'indo-européen commun était donc traversé par des lignes marquant les régions où, à plusieurs points de vue, s'employaient des formes diverses. Et ces lignes, dites lignes d'isoglosses, ne concordent que partiellement les unes avec les autres.

Cet état de choses concorde avec celui que l'on observe en beaucoup de cas. Ainsi le domaine gallo-roman ou le domaine lituanien sont traversés par des lignes d'isoglosses non concordantes entre elles marquant les limites de plusieurs traits linguistiques.

Cette coïncidence entre des types de faits actuellement observables et ceux qu'on est amené à attribuer à l'indo-européen montre combien réelle est l'image que la comparaison amène à se faire de l'indo-européen. Il ne s'agit pas d'un simple schéma, mais d'un objet varié comme la réalité elle-même.

Beaucoup de changements différents suivant les régions naissent après l'époque de communauté, en conséquence des innovations qui sont intervenues. En voici un exemple typique. Le celtique a hérité d'un ancien q^w, du type de celui qui est conservé dans lat. *quis*. Il offre, d'autre part, une tendance qui apparaît dans toutes les parties du groupe où est conservé ce phonème : le k^w tend à passer à *p*, et le g^w à *b*. Le passage de g^w à *b* a abouti aisément, parce que le celtique possédait déjà un *b* qui a attiré l'ancien g^w. Mais, par une particularité singulière, le celtique avait perdu le *p* indo-européen ; dès lors, pour passer à *p*, le q^w devait donner naissance à un *p* qui n'existait plus dans la langue ; ceci a suffi pour empêcher dans le groupe irlandais le q^w de passer à *p*, et c'est ainsi que **eqwos* « cheval » a abouti à irlandais **eq(os)*, d'où *ech*. En gaulois et en brittonique, au contraire, la tendance de q^w vers *p* a été la plus forte, et q^w a passé à *p*, si bien que le mot **eqwos* est devenu *epos* en gaulois (et une forme pareille en brittonique). La division dialectale s'est produite en vertu d'événements qui ont chance d'être postérieurs à l'époque celtique commune. Une fois produite, elle commande des différences ultérieures ; car les mots présentent dès lors des différences profondes et ont des histoires nécessairement différentes.

La notion de dialecte a, on le voit, quelque chose de flou à tous égards. Le comparatiste ne peut donc s'y tenir ; il est obligé de chercher un terrain d'observation plus exactement délimité.

VI.

LA GÉOGRAPHIE LINGUISTIQUE.

Depuis qu'on s'attache à poser des correspondances précises et régulières entre les faits phonétiques ou morphologiques d'une « langue commune » et des langues ultérieurement attestées, on cherche des types linguistiques purs et uns où les règles s'appliqueraient avec rigueur. Les grandes langues renferment des éléments trop divers. Par définition même les dialectes n'ont pas d'unité. On a pensé que le parler populaire, observé dans une localité de petite étendue, fournirait cette unité élémentaire dont le linguiste a besoin. Et l'on a étudié des parlers locaux.

Il a été fait, de parlers très divers, des monographies, les unes brèves et superficielles, les autres détaillées et approfondies. Les monographies précises apportent des données importantes à qui veut connaître à fond un type de langues. Elles sont hautement instructives pour la linguistique générale. Mais, pour le comparatiste qui veut faire l'histoire d'un groupe linguistique, les monographies de parlers locaux ne suffisent pas et sont malaisées à utiliser.

Sur un domaine d'étendue moyenne comme le domaine gallo-roman, il se rencontre plus de trente mille localités dont il y aurait lieu de décrire le parler. La tâche excède manifestement les forces de travail et les ressources matérielles dont disposent les linguistes. Qu'on la suppose accomplie, le linguiste serait accablé par la masse de ces données qu'il ne saurait manier. Il y aurait d'ailleurs des répétitions infinies : bien que chaque localité ait ses particularités propres, les mêmes faits et types de faits se rencontrent sur des domaines étendus.

Autre difficulté : les données recueillies par des observateurs divers ne sont pas exactement comparables entre elles. Si l'enquête n'a pas été organisée, faite suivant des règles uniques pour toute la région étudiée, chaque monographie se présente d'une manière diverse, et les faits qu'on rencontre dans l'une ne se laissent pas comparer immédiatement à ceux d'une autre. Même dans le cas favorable — et rare — où l'enquête a été organisée et où les monographies sont faites sur un modèle unique, les observateurs n'ont pu ni observer ni noter tout à fait de la même manière. Il est de plus inévitable qu'on trouve des enquêteurs pour certaines régions plus que pour d'autres. Et la densité des observations n'est pas comparable.

Enfin les parlers n'ont pas l'unité qu'on a été porté à leur attribuer *a priori*. Les sujets d'un même village, même petit, offrent souvent des manières de parler diverses suivant l'âge, la condition sociale, les occupations, etc. Tous les sujets ne sont pas indigènes ; tous ne sont pas également fidèles à l'usage local. Si une monographie d'un parler local tient compte de ces différences individuelles, elle devient compliquée, et il est malaisé d'en tirer parti pour la comparaison. Si elle les néglige, elle ne donne pas une idée juste de l'état du parler ; elle simplifie arbitrairement ; elle schématise plutôt qu'elle ne décrit.

Quand on veut procéder à l'étude d'un ensemble de parlers modernes par voie comparative, l'enquête doit être organisée de telle sorte qu'elle puisse servir immédiatement à la comparaison.

Il faut tout d'abord avoir des observations réparties d'une manière à peu près égale sur l'ensemble du domaine étudié. L'idéal serait d'observer toutes les localités. Mais, dans un domaine normal où des parlers semblables se rencontrent sur une surface de quelque étendue et où le parler de chaque localité ne diffère pas essentiellement de celui de la localité voisine, il suffit d'examiner des localités prises au hasard, de manière que l'on ait un réseau d'observations comprenant tout le pays et fournissant des spécimens de tous les types. Plus le réseau est serré, moins on risque de laisser échapper des particularités importantes, plus on est sûr de tracer des limites exactes pour chaque fait. Mais l'essentiel pour le comparatiste est de posséder des données qui permettent d'apprécier d'une manière égale l'ensemble du domaine.

En second lieu, il faut que les données soient comparables les unes aux autres, et, pour cela, que les renseignements recueillis concernent des faits de même ordre : les mêmes mots de façon que l'on ait les formes prises par un mot de la langue commune sur tout le domaine, ou les mots de même sens, les mêmes formes ou des formes grammaticales de même valeur, etc.

Pour répondre à cette double nécessité, il faut dresser un questionnaire qui est rempli dans toutes les localités où doit se faire l'enquête. On note la façon dont des phrases données se disent dans chacune des localités enquêtées. Ce procédé du questionnaire a de graves inconvénients : la langue dans laquelle la question est posée — qui est en principe la langue générale du pays — risque d'influencer le sujet local et de le détourner de son parler propre. Pour avoir une réponse, et une seule, par localité, on est amené à interroger un sujet unique; et, puisque le parler local n'est pas un, ce sujet est plus ou moins impropre à représenter l'ensemble du parler. Le procédé est grossier, approximatif. Mais il est le seul possible.

On peut envisager deux manières de faire l'enquête. Ou bien l'on remet un questionnaire qui est rempli sur place par une personne aussi qualifiée que possible pour indiquer la façon dont les choses se disent dans le parler; c'est ainsi qu'on a procédé en Allemagne. Ou bien l'on envoie un enquêteur qui, sur chaque point examiné, interroge un sujet un et note lui-même la réponse ; c'est ce qui a été fait pour l'atlas des parlers gallo-romans. Muni d'un questionnaire préparé par M. Gilliéron, M. Edmont a visité seul toutes les localités qui devaient être examinées, il a choisi un sujet unique dans chacune et a noté la façon dont ce sujet rendait les phrases du questionnaire. L'avantage de ce second procédé, c'est qu'il offre des témoignages rigoureusement comparables entre eux, et que l'on n'a pas à tenir compte des déformations qui résultent de la personnalité d'enquêteurs divers. Cet avantage est si grand que, pour les enquêtes faites par la suite, on a recouru à ce même moyen, et que, par exemple, pour l'atlas des parlers romans (autres que les parlers du type français) de la Suisse et des parlers italiens du Nord qu'ont organisé MM. Jaberg et Jud, un seul observateur, M. Scheuermeier fait toute l'enquête.

A première vue, les résultats obtenus par des procédés aussi sommaires pourraient paraître suspects. Mais les observations faites sur des points peu éloignés les uns des autres se contrôlent mutuellement, et la vérification résulte du rapprochement même. Du reste, le procédé de l'enquête n'exclut pas celui de la monographie ; et il suffit de comparer les résultats de l'enquête avec les monographies déjà faites ou avec des recherches nouvelles pour apprécier la valeur des observations recueillies par l'enquêteur. En ce qui concerne l'atlas des parlers gallo-romans exécuté par M. Gilliéron avec les observations qu'avait recueillies M. Edmont, il y a eu des vérifications nombreuses : en plusieurs régions, il a été procédé à des enquêtes détaillées portant sur tous les parlers d'une région limitée, ainsi M. Millardet pour les Landes, M. Bruneau pour les Ardennes, M. Oscar Bloch pour les Vosges ; ces enquêtes minutieuses ont en général confirmé pour l'essentiel la justesse des observations de M. Edmont. On a pu relever des erreurs de détail ; nulle part, ces menues erreurs ne vicient l'enquête d'une manière fondamentale.

Les résultats obtenus ont été capitaux, et tels qu'on s'est mis à multiplier les enquêtes de ce genre. En France, il a été fait un atlas pour la Corse, et un autre paraît pour les parlers bretons d'Armorique française. Un atlas catalan est en voie de publication. On a signalé déjà l'atlas que préparent MM. Jaberg, Jud et Scheuermeier. Un jeune Français, M. Tesnière, publie actuellement, sur les formes du duel en slovène, une étude détaillée fondée sur une enquête du type géographique.

Ceci seul que les résultats de l'enquête se prêtent à être indiqués sur des cartes, consacrées chacune à un fait donné, facilite le travail. On sait ce qu'une statistique gagne en clarté à être exprimée par un graphique. Le linguiste qui voit sur une carte unique, ou sur deux ou trois cartes qu'il peut juxtaposer, les faits relatifs à un problème a, d'un coup, sous les yeux, les éléments essentiels de la solution.

Dès maintenant, les recherches sur le gallo-roman ont été renouvelées par l'*Atlas* de MM. Gilliéron et Edmont. Les études provoquées par la publication de l'*Atlas* se multiplient, et d'autres études où il est tenu compte avant tout des données recueillies sur tout l'ensemble d'une région paraissent sans cesse.

C'est que la comparaison a trouvé, dans ces enquêtes, un instrument de travail supérieur à tout ce qu'elle possédait et précisément adapté à ses besoins. Pour la première fois, on avait, clairement présenté, un ensemble de données immédiatement comparables entre elles, et réparties sur tout le domaine étudié. Quiconque a fait des travaux de grammaire comparée sait combien on souffre de ce que les faits rapprochés offrent des différences de niveau dont il faut faire abstraction : le comparatiste qui travaille sur les langues indo-européennes se sert de données dont les dates s'étendent sur un espace de quelque trois mille ans, qui abondent à certains moments et manquent tout à fait à d'autres, qui existent pour une région alors que, pour tel autre domaine, toute indication manque. Avant de faire un rapprochement, il en faut critiquer en détail tous les éléments. Dans la grammaire comparée des langues indo-européennes, il y a peu de rapprochements qui ne boitent pas par quelque côté.

Dès maintenant, les résultats obtenus par la méthode géographique sont saisissants.

Le problème du « dialecte » a trouvé sa solution. On s'était souvent demandé comment tracer les limites entre les dialectes. D'une part, le dialecte apparaissait comme un ensemble offrant des caractères particuliers et s'opposant à d'autres dialectes. De l'autre, on n'arrivait pas à trouver aux dialectes des limites précises. Autrefois, un observateur parti de Bordeaux pour déterminer la limite entre les parlers gallo-romans du Nord et du Midi, avait dû renoncer à marquer une frontière et s'était arrêté sans avoir terminé son travail. Il a suffi de rapprocher les cartes pour apercevoir la vérité.

Chaque fait linguistique a ses limites propres.

Mais, en certaines régions, on rencontre un faisceau de limites de faits particuliers, limites qui parfois sont parallèles, parfois se croisent. Entre les parlers gallo-romans du Nord et ceux du Midi, il n'y a pas une limite ; il y a un faisceau de limites de faits particuliers, limites dont beaucoup sont proches les unes des autres. Au Nord et au Sud de ce faisceau, les parlers offrent par suite de nombreux traits en commun. L'opposition du Nord et du Sud est nette ; mais elle ne se laisse pas marquer par une ligne unique.

Ce faisceau suit, d'une façon caractéristique, à peu près le chemin que suit encore le train du chemin de fer qui va de Bordeaux à Lyon, c'est-à-dire qu'il fait une pointe prononcée vers le Nord, pour passer au pied du massif central de la France, du côté Nord. Tout le massif central appartient donc au type des parlers méridionaux. Les influences qui ont donné aux parlers de la partie septentrionale de la France leur aspect particulier, si éloigné du type latin, se sont arrêtées devant le massif montagneux qui occupe le centre de la France.

Ce n'est pas la Loire qui est une limite; c'est un groupe de hauteurs, en effet pénible à traverser et que les routes contournent nécessairement. Le type français du Nord va beaucoup plus au Sud, à la fois à l'Ouest, où il atteint Bordeaux, et à l'Est, où il va au Sud de Lyon. Ici la limite des dialectes s'explique par un fait d'ordre géographique, qui a conditionné des actions historiques.

En second lieu, grâce à la grande multiplicité des faits observés, la période intermédiaire entre la langue commune et les langues ultérieures s'éclaire. C'est le nom de l'abeille que M. Gilliéron a choisi pour illustrer, d'une manière particulièrement approfondie, les résultats de la méthode géographique. On savait, par la considération des règles de correspondance phonétique, que le mot français *abeille* ne peut passer pour indigène dans la France du Nord et qu'il est emprunté à des parlers méridionaux. Mais ce fait brut ne révélerait pas l'histoire du mot *abeille* en français. Il a suffi d'examiner la carte où sont pointés les noms de l' « abeille » pour apercevoir que ces noms varient suivant les régions du Nord de la France : ici *abeille,* terme venu du Midi ; là *mouche à miel,* là *mouchette,* là *avette.* Cette variété, pour un insecte utilisé partout depuis l'époque romaine, montre que le nom ancien a disparu ; or, les cartes révèlent, sur quelques points, un nom éliminé : c'est *é,* c'est-à-dire la forme que le latin *apem* a prise normalement dans la France du Nord. Trop bref, ce nom a été presque partout éliminé pour être remplacé par des noms nouveaux qui marquent l'embarras où se sont trouvés les sujets parlants. Dans le mot *mouchette,* M. Gilliéron trouve, par une hypothèse hardie et qui a été contestée mais qui est singulièrement séduisante, une adaptation d'un composé *mouche-é(p)* auquel on aurait recouru pour se tirer de l'embarras causé par la trop grande réduction de *apem* dans la France du Nord ; pareil composé a pu d'ailleurs n'exister que virtuellement, et *mouchette* a pu s'y substituer de manière immédiate ; car, au moment où se produisent les changements de ce genre, phonétiques ou morphologiques, les innovations tendent à se régler aussitôt d'après le système existant et à prendre place parmi les formes déjà constituées. Le Midi, qui avait *apicula*, élargi dès l'époque romane, tout comme *ouis* a été élargi en *ouicula* (fr. *ouaille*), n'a pas connu pareille difficulté et a fourni *abeille* à la région parisienne. A l'indication schématique qu'on avait sur l'origine de fr. *abeille* se substitue une histoire complexe et où apparaît la réalité.

En s'inspirant de ces résultats, on interprétera mieux les faits là même où l'on dispose de données moindres et moins aisément comparables. Bien souvent il suffit de disposer les faits géographiquement pour en comprendre l'histoire.

Soit, par exemple, en indo-européen la notion de « moudre ». Il y a un présent du type **meləmi* ou **moləmi* « je mouds » qui est conservé indirectement en slave, en baltique, en germanique, en celtique et en italique et dont on a pu restituer ci-dessus la forme ancienne (v. p. 28). Or, ce verbe ne se retrouve ni en indo-iranien, ni en arménien ni en grec : ce n'est pas un accident, et ce n'est pas que les peuples qui parlaient ces langues aient ignoré la mouture. Car des noms de céréales figurent justement dans ces langues ; on y a (sauf en indo-iranien) la racine signifiant « labourer ». Mais c'est qu il y a une autre racine pour signifier « moudre » :

aléō « je mouds » en grec, *ałam* en arménien, et cette racine a laissé des traces nettes aussi en indo-iranien. On voit donc que l'idée de « moudre » s'exprime de façons différentes sur deux domaines continus de l'indo-européen.

Quant à la racine **melə-* qui signifie « moudre », du slave à l'italique, elle n'est pas inconnue aux autres langues ; mais elle s'y trouve au sens plus général d' « écraser, broyer », ainsi dans *mṛṇā́ti* « il écrase » du sanskrit, dans *malem* « je broie » de l'arménien ou dans des verbes du grec populaire tels que *myllō* ou la forme à redoublement *moimyllō*, qui se lit chez des auteurs ioniens populaires. Et la concurrence entre la racine **melə-* et celle du grec *aleō* ressort du fait que, de même que le latin a *mola* « meule » à côté de *molo* « je mouds », le grec qui a perdu le nom de l'ancienne meule à bras (conservé depuis l'Inde jusqu'au monde celtique), a formé *mylē* « meule », de la racine signifiant « écraser ». La répartition géographique des mots laisse entrevoir toute une histoire, malgré la misère des données.

Du rapprochement des cartes ressort souvent l'explication de faits surprenants. Ainsi, il y a au Sud de la France une région où le représentant attendu du nom *gallu(m)* du « coq » a disparu. Dans cette petite région, on désigne le « coq » par des noms bizarres qui traduisent un embarras où s'est trouvée la langue : le « coq » est nommé par exemple « faisan » ou « vicaire ». Mais si l'on rapproche d'autres cartes, on voit que, dans cette région, c'est *-tt-* qui répond à *-ll-* du latin : le nom **gat* du « coq » se confondait dès lors avec le nom *gat* du « chat », ancien **gattu(m)*, qui a donné aussi *gatto* de l'italien. Il serait incommode de désigner par le même nom le « chat » et le « coq » ; le nom du « coq » a disparu. Et la langue s'est tirée d'affaire par des artifices visibles. Quand une notion vient à être indiquée par des procédés aussi singuliers, c'est en général qu'un nom ancien, exclu pour une raison ou pour une autre (les causes sont diverses), a dû être remplacé ; la fantaisie des sujets parlants se donne cours en pareil cas.

Partout où l'on a pu appliquer la méthode géographique, elle a donné lieu à des progrès décisifs. Elle exige des enquêtes aussi étendues qu'il est possible et l'utilisation de toutes les données qu'on possède sur l'ensemble d'un domaine linguistique. La méthode comparative gagne par là une précision, une étendue et une aisance jusqu'alors imprévues.

Et c'est pourtant ce que l'on devait attendre : la grammaire comparée, qui opère avec des systèmes de correspondances, gagne à considérer moins les parlers — qui ne l'intéressent pas par eux-mêmes — que des faits d'un même ordre dans l'ensemble des parlers qui continuent une même « langue commune ».

Il est inutile de rappeler que, comme on l'a noté déjà, p. 29, tout parler ayant son système propre, il faut toujours se représenter la place de chacun des faits de détail dans chacun de ces systèmes. Un examen exclusif des mots et des formes fournis par les enquêtes et notés géographiquement risquerait d'entraîner à étudier d'une manière isolée un mot ou un petit groupe de mots, une forme ou un petit groupe de formes. Pareil émiettement ruinerait toute linguistique historique. La géographie linguistique a eu le mérite de mettre en pleine évidence la singularité de l'histoire de chaque mot, de chaque forme. Mais cette singularité a sa place dans des ensembles systématiques, et qui envisagerait les faits isolés sans se les représenter dans ces ensembles risquerait de commettre des erreurs pires encore que le linguiste qui envisage exclusivement des ensembles et n'étudie pas avec une critique assez sûre chacun des faits particuliers dont sont faits ces ensembles.

VII.

LA NOTION DE LANGUE MIXTE.

Durant longtemps, la linguistique historique a opéré avec l'hypothèse tacitement acceptée que la langue se transmet de génération en génération, chaque enfant reproduisant de son mieux le parler de son entourage. Tel est le cas qui semble normal et qu'on observe souvent en effet.

Mais il arrive qu'une population change de langue. Ce n'est pas un fait exceptionnel : parmi les pays dont l'histoire est connue — et l'on sait que nulle part, dans le cas le plus favorable, l'histoire faite au moyen de textes ne remonte au delà de cinq mille ans environ, période courte par rapport au développement total de l'humanité —, il n'y a presque pas un peuple qui n'ait changé de langue au moins une fois, et généralement plus d'une fois.

De tous les pays, l'Égypte est celui dont l'histoire remonte le plus haut et où il y a le plus de conservation ; or, après avoir persisté durant environ quatre mille ans de période historique, l'égyptien est sorti de l'usage et a été remplacé par l'arabe, sans que le fond de la population ait sensiblement changé ; la conquête macédonienne a amené une colonie grecque, surtout à Alexandrie, et elle a entraîné l'hellénisation de petites couches de dirigeants ; mais le fond de la population est demeuré le même, même en tenant compte des émigrants venus de Syrie. — Sur le territoire de la France actuelle, le gaulois a dû arriver avec la conquête celtique, durant la première moitié du millénaire qui a précédé l'ère chrétienne ; puis il a cédé la place au latin après la conquête romaine ; la conquête gauloise ne paraît pas avoir renouvelé la population, et la conquête romaine n'a sûrement apporté que peu d'éléments nouveaux. — Il serait aisé de multiplier les exemples : il intervient souvent des changements de langue. Il importe d'examiner en quelles conditions se font ces changements et quels en sont les effets.

Le cas le plus facile à observer est celui qui se rencontre actuellement presque partout en Europe. Dans chaque région, il y a un groupe de parlers locaux de même famille et une langue écrite, langue de civilisation qui sert à tous les usages généraux, aux relations avec l'ensemble du pays et qui est la langue du gouvernement, de l'école, des administrations, de la presse, etc. En pareil cas, la langue écrite a sur les parlers locaux une forte influence.

Alors il se produit deux mouvements de sens divers et qui aboutissent à un même résultat final.

Le parler local s'emplit d'éléments empruntés à la langue générale, qui seule est en mesure de répondre aux besoins nouveaux de la civilisation et dont le prestige est grand par rapport au parler local : c'est la langue des gens les plus puissants et les plus cultivés. En étudiant les parlers locaux de la Gaule romane,

M. Gilliéron a reconnu que, dans les cas où ils éprouvaient un embarras, ces parlers ont cherché dans le français général les moyens de se tirer de difficulté. Ainsi, loin de fournir une tradition pure et résultant de leur histoire propre, les parlers gallo-romans comprennent, en une large mesure, des éléments français généraux « patoisés ». Les études sur les parlers gallo-romans comportent maintenant un grand chapitre sur les procédés par lesquels ces parlers ont adapté les mots français. Comme le français et le parler sont de même famille, ils présentent pour le fonds ancien de leur vocabulaire des correspondances régulières, et, en tenant compte de ces correspondances, il n'est pas difficile aux sujets employant le parler local de faire, avec des mots français, des formes de type patois, ou de transporter dans le patois des tours de phrase français. En revanche, ce travail d'illettrés est souvent si bien fait que le linguiste a peine à discerner dans le patois ce qui est vraiment indigène de ce qui est adapté.

La supériorité de la langue générale et par sa valeur et par les commodités qu'elle présente pour communiquer avec d'autres hommes est telle que, en dépit des vieux usages, on tend bientôt à préférer cette langue générale au parler local, comme on substitue le costume urbain aux costumes locaux. Mais il est moins facile de changer de langue que de costume. Si les parlers locaux de la France sont souvent en grande partie du français patoisé, le français employé au village est souvent du patois francisé. Dans bien des cas, il suffit de changer un peu le mot local pour en faire un mot français correct : le mot *lwé* du village devient français dès qu'on le prononce *lwa* (*loi*). Le mot berrichon *lavwé* est un peu plus difficile à franciser ; il ne suffit pas de le prononcer *lavwa* : il faut ajouter *r* (fr. *lavoir*); aussi y a-t-il chance que *lavwé* survive quelque temps ; mais tout le monde finit par dire *lavoir*, à la manière française. Ce qui survit le plus, ce sont certains usages grammaticaux fixés dans le mémoire et dont on prend à peine conscience, comme il a été indiqué ci-dessus, p. 26. Ces usages s'éliminent à leur tour, et ainsi, morceau par morceau, le parler local est remplacé par le français général. Ce mouvement se produit maintenant d'une manière rapide, plus vite dans certaines régions que dans d'autres : le français général élimine plus le parler local dans le centre de la France qu'en Picardie par exemple. Mais le mouvement a lieu partout, et dès maintenant on entrevoit le moment où, dans la France du Nord, toute étude des parlers locaux sera devenue impossible, parce que ces parlers seront sortis de l'usage. Or, ils étaient encore usuels partout dans les villages à la fin du XVIII^e^ siècle ; et, encore dans la première partie du XIX^e^ siècle, la façon de parler dans tous les villages de France conservait au moins des restes considérables du type local.

Le fait français n'est pas une nouveauté dans l'histoire. Au V^e^ siècle av. J.-C., presque chaque localité grecque avait son parler propre. A partir de cette époque, l'action de la langue générale, de plus en plus forte, élimine les unes après les autres les particularités locales, et une langue commune, fondée sur l'usage attique, se répand sur toute la Grèce. Malgré des réactions multiples, cette langue s'est imposée partout, et, dès l'époque impériale, on voit que l'emploi en était devenu à peu près universel sur le domaine hellénique. Les parlers grecs modernes, qui ont de nouveau divergé les uns d'avec les autres, continuent tous cette langue commune. Et c'est à peine si l'on entrevoit sur un point quelque menue survivance des parlers locaux du grec antérieurs à la généralisation de la langue commune.

Mais cette généralisation ne va pas sans des réactions qui modifient la langue générale. Le grec qui s'est généralisé n'est pas du pur attique. Les traits qui se sont répandus partout sont surtout ceux qui étaient communs à l'attique et à l'ionien et, en outre, quelques particularités propres à l'attique. Mais d'autres particularités trop spéciales à l'attique n'ont pas été acceptées. Par exemple on a préféré *ss* à la prononciation *tt* des formes correspondantes de l'attique, et l'on a dit *melissa* « abeille », comme en ionien et dans la plupart des parlers, non pas *melitta,* qui était la forme d'Athènes. Comme, avec le temps, Athènes s'est conformée aussi à l'usage commun, des particularités attiques ont disparu. Ainsi « aucun » était arrivé au ɪvᵉ siècle av. J.-C. à se dire *uthēs* à Athènes ; cette forme attique n'a pas prévalu, et la vieille forme *udēs,* qui avait subsisté hors d'Athènes, a fini par revenir à Athènes. Le grec moderne, pour dire « non », a en conséquence *ðen,* qui est l'ancien neutre *udén* « rien ». Le grec commun ne continue donc pas purement un seul parler ; il comprend des éléments appartenant à des parlers voisins les uns des autres, mais distincts.

Même le français général, qui est la langue de Paris, n'est pas sans subir des influences provinciales, malgré le purisme qui prévaut en France.

Ainsi une langue générale est plus ou moins une langue mixte, par la manière dont elle se constitue et par la complexité des éléments qui y entrent.

Ceci peut aller loin. L'italien général est fondé sur l'usage toscan; mais la prononciation n'admet pas les particularités toscanes, et l'on enseigne volontiers que l'italien général est une *lingua toscana in bocca romana.* L'anglais général repose sur l'usage de Londres; mais, placé au point de rencontre de dialectes divers, Londres a dans son parler des traits qui proviennent de régions différentes. Ainsi les langues générales, par le fait qu'elles se sont fixées pour répondre aux besoins d'hommes dont les origines sont différentes, ne reposent pas sur une tradition une, même dans les cas les plus favorables.

Il peut même y avoir de l'artifice dans la manière dont se constituent les langues générales. Ainsi le grammairien Vuk qui a donné au serbo-croate écrit sa forme, s'est fondé sur les parlers d'une région serbo-croate où l'ancien *x* (*h*) était amui; et il avait commencé par fixer la langue sans *h*. Après réflexion, il lui a paru bon de tenir compte des parlers, de type plus archaïque, où *h* s'était conservé. La langue qu'il a fixée s'est trouvée ainsi être composite. Et, comme les sujets appartenant à des régions où *h* avait disparu avaient peine à prononcer *h*, il y a sur ce point un flottement dans l'usage serbo-croate général et beaucoup de trouble.

L'extension d'une langue générale n'est pas limitée aux cas où les parlers locaux appartiennent au même type linguistique que cette langue générale et ont une même origine. Quand la langue générale est nettement différente des parlers locaux, les faits se présentent d'une manière autre que dans le cas étudié; mais le développement est semblable à plusieurs égards.

D'une part, la langue indigène à laquelle se superpose une langue générale différente, se réduit à des usages locaux et inférieurs. Dès lors, elle se charge d'éléments empruntés à la langue générale. C'est ainsi que, dans l'Armorique française, le breton s'emplit de mots et de tours français. Il peut être parlé avec négligence et perdre quelques-unes de ses particularités les plus caractéristiques. Par exemple, le breton de la région de Vannes s'est beaucoup détérioré. Une langue réduite à des usages locaux tend à perdre une partie de plus en plus grande de ses éléments. En général, ceci importe peu pour l'histoire de la langue ; car un idiome aussi dégradé finit par s'éliminer tout à fait.

L'usage de la langue générale, plus commode, et qui classe le sujet parmi les gens cultivés, se répand de plus en plus. Mais les gens qui passent ainsi à la langue générale ne sauraient la parler exactement comme ceux pour qui elle est traditionnelle. Outre des termes locaux désignant des animaux, des plantes, des usages, des formes de pays pour lesquels la langue générale n'a pas de mots aussi précis, il subsiste chez les sujets qui sont passés à une langue nouvelle des manières anciennes de s'exprimer ; la prononciation n'est d'ordinaire pas reproduite d'une manière parfaite.

Qu'on observe, par exemple, ce qui se passe en France. Bien que les parlers du Nord et du Midi soient également gallo-romans, ils diffèrent assez pour qu'on ne se comprenne absolument pas, avec le parler local, d'une région à l'autre. Quand les méridionaux emploient le français général, qui, pour les habitants des villes, est souvent aujourd'hui la seule langue qu'ils connaissent, ils le prononcent d'une manière spéciale, et qui d'ailleurs diffère d'une région à une autre. La caractéristique des voyelles est plus haute dans le Midi de la France que dans le Nord ; les timbres vocaliques employés sont autres ; les voyelles nasales sont autrement articulées ; les *e* muets, dont la répartition en français est chose délicate et subtile, sont placés de manière incorrecte ; bref, l'ensemble de la prononciation diffère. La grammaire même comporte des différences : le prétérit simple, dont la langue écrite a maintenu l'emploi, mais qui n'existe plus dans le français parlé de la région parisienne, est fréquent chez les méridionaux parlant français. Ainsi, même dans un cas où, comme en français, il y a un type unique d'enseignement de la langue, une influence puriste puissante et un usage fixé avec une rigueur nulle part dépassée, rarement égalée, la généralisation de la langue ne va pas sans une forte adaptation aux habitudes locales. En se généralisant, le français se brise en français régionaux, et le modèle parisien n'est exactement reproduit dans aucune province. Si ces français régionaux n'ont jamais été décrits avec précision, ils n'en existent pas moins. Comme on dit, l'influence du « substrat local » se fait sentir, et, pour autant qu'on peut observer l'extension des langues générales, cette influence paraît inévitable.

Sans doute l'action de l'école et de la langue écrite et le désir qu'ont les sujets de parler correctement tendent à unifier la forme de la langue générale ; mais cette unification ne saurait être parfaite. La norme de la langue peut être la même partout ; mais elle ne se réalise pas de même partout.

Sans doute même y a-t-il quelque chose de plus profond : des expériences indiquent que, si les caractères anatomiques acquis ne se transmettent pas, il y aurait une hérédité des habitudes acquises. Or, l'usage linguistique a, au plus haut degré, le caractère d'un ensemble d'habitudes acquises. En apprenant une langue nouvelle, les sujets ne perdent pas pour cela leur hérédité. Des tendances qui se font jour dans la langue indigène peuvent dès lors se manifester aussi dans la langue nouvellement adoptée. On comprend ainsi que le latin ait subi des transformations particulièrement profondes sur les domaines où il a remplacé d'autres langues à tendances toutes différentes, et notamment le gaulois. On sait, par exemple, que l'un des traits qui caractérisent le celtique est une forte tendance à altérer, à conformer aux voyelles voisines et même à éliminer les consonnes placées entre voyelles. Or, nulle part plus que sur le domaine gallo-roman, et surtout en français, les consonnes intervocaliques latines ne se sont altérées ou même n'ont été éliminées comme elles l'ont été en français : *lepore(m)* devient *lièvre*, et *amatam* devient *aimée*. On est amené ainsi à supposer que les innovations les plus caractéristiques du français tiendraient, non seulement à la façon dont le latin a été prononcé en Gaule, mais à une hérédité d'habitudes acquises par les sujets parlant gaulois. Avec cette forme de l'hypothèse, les objections que l'on fait souvent à la théorie du *substrat* se résolvent immédiatement. Le passage de *u* fermé à *ü* sur le domaine gallo-roman et en Alsace ne serait pas une survivance immédiate du gaulois, mais l'effet lointain de certaines habitudes acquises transmises par hérédité.

Si l'explication — au moins partielle — de la diversité des formes prises sur différents domaines par une langue qui se généralise est souvent contestée, c'est qu'on envisage les faits d'une manière trop mécanique : il ne faut pas s'attendre à trouver dans la langue généralisée des particularités du « substrat » sous forme brute. L'action est complexe et se manifeste sous des formes au premier abord peu apparentes.

Ceci posé, une langue généralisée est, sur les domaines où elle s'étend, une langue plus ou moins mixte, puisqu'à côté d'une norme universelle, on y observe l'influence — sans doute profonde — d'usages locaux, d'habitudes locales.

D'ailleurs, et même en dehors de ces influences intimes, l'emprunt suffit à donner à une langue un aspect nouveau s'il devient fréquent. Le fait que l'anglais doit au latin et au français plus de la moitié de son vocabulaire suffit à lui seul à le différencier d'avec les autres langues germaniques. La prononciation même en est touchée, puisque les mots empruntés ont gardé leur accentuation et que l'accentuation germanique sur l'initiale du mot a cessé, dès lors, de garder son caractère de constance. On peut emprunter des manières de parler : le tour allemand *was für* peut être transposé dans un tour slave *čto za*. Il se produit même des expressions dont les éléments sont pris à deux langues : en latin, le grec *kath-ena* (littéralement « par un ») a été en partie calqué, en partie emprunté, et il est résulté de là **cata-uno, *cat-uno,* d'où espagnol *cadauno,* italien *caduno* « chacun ».

Qui dit emprunt admet que le sujet parlant a conscience de deux langues distinctes. Et en effet, dans les langues à morphologie compliquée comme les langues indo-européennes, les langues sémitiques, les langues bantous, ou même dans des langues à morphologie plus simple, mais comportant des caractéristiques bien définies de catégories grammaticales, le sujet parlant doit toujours parler une certaine langue; il a le sentiment de parler ou une langue ou une autre; et il ne saurait mêler la morphologie d'une langue à celle d'une autre. Si dégénéré que puisse être son breton, l'habitant de la région de Vannes sait s'il parle français ou breton. Car il emploie suivant les cas deux morphologies essentiellement distinctes. M. Marr a supposé que la morphologie arménienne, dont le caractère indo-européen est manifeste, comprendrait des éléments dus à une population antérieure à l'arrivée dans le pays arménien d'hommes parlant une langue indo-européenne; mais il n'a pas réussi à démontrer cette hypothèse et il n'a, hors de son école propre, convaincu aucun des linguistes qui s'intéressent à l'histoire de l'arménien.

Pour des langues comme celles d'Extrême-Orient qui ont un minimum de morphologie, et où la phrase se fait à l'aide d'ordres de mots définis et de mots accessoires, on conçoit mieux que le sujet parlant emprunte des procédés à deux langues. On se représente aisément qu'un Annamite puisse mêler à son parler propre des traits du chinois. La juxtaposition de deux langues de cette sorte est relativement aisée. Et ce n'est peut-être pas un hasard qu'il soit difficile, en Extrême-Orient, de reconnaître et d'établir des parentés de langues.

En fait il ne s'est pas rencontré jusqu'ici de cas où l'on ait été conduit à poser que le système morphologique d'une langue donnée résulte d'un mélange des morphologies de deux langues distinctes. Dans tous les cas observés jusqu'à présent, il y a une tradition continue d'une langue ; il peut y avoir un nombre plus ou moins grand d'emprunts, et ces emprunts peuvent occasionner certains procédés nouveaux de formation des mots ; la tradition peut être du type courant : transmission de la langue des anciens aux jeunes, ou résulter d'un changement de langue : apprentissage d'une langue de communication par une population qui abandonne finalement son idiome propre. Mais il y a tradition une pour le système morphologique.

Néanmoins on ne saurait affirmer que, dans certains cas favorables, il n'y ait pas eu des « mélanges » véritables. Le jour où l'on viendrait à les rencontrer, la tâche du linguiste serait malaisée. Si l'on a pu arriver à faire, par la comparaison, l'histoire de quelques langues, c'est que l'on était sûr que chaque système nouveau devait s'expliquer en partant d'un système unique. Au cas où l'on devrait faire état de deux systèmes initiaux et de réactions de l'un sur l'autre, les méthodes actuelles ne suffiraient pas. Car le droit qu'on aurait de choisir entre deux séries de formes originelles causerait un tel arbitraire que toute démonstration deviendrait presque irréalisable. Malgré les hypothèses faites en ce sens, les linguistes ne se sont par bonheur jamais jusqu'ici trouvés de manière sûre devant pareille difficulté. Pour en triompher, la linguistique devra, si la difficulté vient à se rencontrer vraiment, élaborer de nouvelles méthodes, plus délicates que celles qui sont décrites ici, et il resterait à les éprouver.

VIII.

LES FORMULES GÉNÉRALES DE CHANGEMENT

Tant pour répondre à son usage qui est de servir de moyen de communication entre membres d'une société que parce que la tendance normale de tout groupe social est d'établir une exacte conformité entre ses membres, la langue de chaque groupe social tend à être une. La manière de prononcer, la morphologie dont se servent les sujets parlant une même langue concordent à beaucoup d'égards avec une précision rigoureuse. Des détails de formes, des nuances de sens que le grammairien le plus exact a peine à discerner et à formuler se maintiennent souvent pendant des suites de générations sans subir aucune altération notable. Ces concordances minutieuses n'excluent pas de fortes divergences portant sur quelques points ; mais ces divergences partielles ne sont possibles que justement parce que la langue a un fond d'unité, parce que les sujets parlants d'un même groupe sentent qu'ils sont d'un même type.

Quand des changements viennent à se produire, ils doivent dès lors porter sur le groupe dans son ensemble, sous peine d'en entraîner la dislocation. Et l'on constate en effet que les changements linguistiques atteignent l'ensemble des sujets. Le langage ne vit que de communauté; si les sujets se mettaient à modifier la langue d'une manière indépendante, elle ne répondrait plus à son objet. Comme les conservations des usages anciens, les innovations sont donc générales. Sur la façon dont s'établit cette généralité, on n'a pas encore de données certaines. On aperçoit au moins que les innovations sont en partie générales et en partie généralisées. Elles dépendent de conditions communes à tous les sujets, mais la fixation est due sans doute pour beaucoup au sentiment qu'ont les sujets de devoir employer un même type linguistique. La société réagit souvent d'une manière forte contre les innovations individuelles en matière de langue.

Ce qui montre que le changement ne tient pas à la généralisation plus ou moins fortuite d'innovations individuelles, c'est que des changements parallèles ont lieu dans des cas dont le linguiste voit la similitude, mais où le sujet parlant ne peut apercevoir que des différences. Les langues romanes offrent à cet égard un exemple significatif. L'ancien *ō* long et l'ancien *ŭ* bref du latin ont abouti à un *o* fermé comme la forme ancienne de la plupart des parlers romans; ainsi la voyelle accentuée de *nōdum* et celle de *gulam* aboutissent à une même voyelle *eu* dans *nœud* et *gueule* du français (la différence d'orthographe est sans signification, et les nuances qu'on observe tiennent à la différence de position en français); or l'ancien *ē* long et l'ancien *ĭ* bref ont de même abouti à *e* fermé: lat. *mē* a abouti à *moi* et lat. *bibit* à *boit* en français. Les deux confusions s'expliquent par des conditions communes, dont les principales sont la perte des oppositions quantitatives et le fait que, en latin ancien, les voyelles longues étaient fermées et les voyelles brèves ouvertes :

ō s'est trouvé rapproché de *ŭ*, comme *ē* de *ĭ*. Le rapport de *ō* à *ŭ* est le même que celui de *ē* à *ĭ* ; mais, pour le sujet parlant, ce fait théorique n'est pas visible. Le parallélisme des innovations ne peut résulter que de l'effet commun de conditions pareilles.

Si les innovations sont ainsi à la fois communes à divers sujets, sans qu'il y ait action des uns sur les autres, et parallèles entre elles, c'est que dans une certaine mesure, elles dépendent de conditions générales.

Tous les linguistes qui ont eu à examiner des changements phonétiques et à fixer des règles de correspondances entre des langues diverses ont senti que ces changements ont lieu suivant certains types généraux. Quelques types de faits s'observent dans un grand nombre de cas. Par exemple, les occlusives dites gutturales — en réalité caractérisées par le contact ou le rapprochement de la surface de la langue et du palais — sont sujettes à se « palataliser » devant des éléments prépalataux tels que *y* les voyelles et *i*, *e* et même la voyelle *a* orientée vers *e* ; c'est ce que l'on note en marquant que, en pareil cas, *k*, *g* deviennent *k'*, *g'* : le *k* de fr. *qui* est tout autre chose que celui de fr. *cou*. Or, ces *k'*, *g'* prépalataux sont sujets à devenir *tš*, *dž* (*č*, *ǰ*) ou *ts* (*c*), *dz* ; et ces *tč*, *dž* et *ts*, *dz* à devenir simplement *š*, *ž* et *s*, *z*. Par exemple, lat. *cinere*(*m*) est devenu fr. *cendre* et lat. *carbōne*(*m*) fr. *charbon*. Des développements pareils s'observent un peu partout dans le monde et dans les familles de langues les plus variées.

Mais il y a des formules plus délicates dont M. Grammont a reconnu l'existence dans son travail fondamental sur la dissimilation consonantique. Si, dans un même ensemble d'émissions phoniques, un mouvement articulatoire doit être répété, le sujet parlant tend à éviter cette répétition.

Cette tendance n'atteint pas également tous les phonèmes ; il y a des mouvements articulatoires moins faciles et courants que d'autres, et ce sont ceux-ci qu'on vise à ne pas faire deux fois à bref intervalle. La prise des positions particulières que la langue doit occuper pour *l* et *r*, l'abaissement du voile du palais qui est essentiel pour *n* et *m* sont des mouvements dont la répétition est évitée souvent.

Alors, c'est dans un ordre défini que la suppression a lieu. Par exemple, si l'ancien *uenēnu*(*n*) aboutit en italien à *veleno*, c'est-à-dire si, de deux *n* que comprend le mot, l'une a perdu l'abaissement du voile du palais et a, en conséquence, passé à *l*, c'est la première qui est ainsi altérée ; sauf réaction due à des faits particuliers, de deux *n* ainsi placées c'est la première, non la seconde, qui perd sa nasalisation. Ce sont donc des conditions générales qui règlent la façon dont se produit la dissimilation, et ces conditions se laissent formuler. M. Grammont a réussi à les déterminer entièrement.

Ceci permet de reconnaître, en bien des cas, la manière dont une innovation a eu lieu. On sait, par exemple, que, dans les langues où l'accent sert de centre au mot, les voyelles de syllabes accentuées sont traitées autrement que les voyelles de syllabes inaccentuées. Or, on observe ici deux types distincts d'altération des voyelles inaccentuées.

Dans certaines langues, les voyelles inaccentuées tendent à perdre à la fois une partie de leur durée et beaucoup de leur timbre propre : elles se réduisent souvent à un timbre plus ou moins neutre, tel qu'est celui de l'*e* allemand des syllabes inaccentuées ; cet *e* peut, suivant les cas, reposer sur une voyelle ancienne de timbre quelconque. On observe actuellement pareille altération en russe : les voyelles inaccentuées y deviennent de plus en plus indistinctes ; elles ont de moins en moins un timbre propre, tandis que les voyelles accentuées sont prononcées nettement.

Dans d'autres langues au contraire, les voyelles inaccentuées s'altèrent en fonction de leur degré d'ouverture : *e* et *o* deviennent *i* et *u* ; et les voyelles *i* et *u* s'amuissent purement et simplement ; c'est ce qui s'est produit dans les parlers grecs modernes du Nord.

Or, en allemand, en anglais ou en russe, les syllabes accentuées sont avant tout intenses, et beaucoup plus intenses que les inaccentuées ; au contraire, en grec moderne, les voyelles des syllabes accentuées sont caractérisées surtout par leur durée : elles sont plus longues que les autres. Si les voyelles inaccentuées tendent à perdre leur timbre propre, c'est donc en vertu de l'intensité plus grande des voyelles accentuées ; si les voyelles inaccentuées tendent à se fermer, *e* et *o* en *i* et *u*, c'est en vertu de la durée plus grande des voyelles accentuées : *i* et *u*, qui sont déjà par nature des voyelles plus brèves que *e* et *o*, ne peuvent s'abréger davantage qu'en disparaissant.

Dès lors, là où l'on observe des faits comparables à ceux du type grec septentrional, il y a lieu de croire que c'est en vertu d'oppositions quantitatives. Par exemple, rien ne permet d'attribuer à l'accent arménien une intensité comparable à celle de l'accent anglais, allemand ou russe moderne : l'arménien actuel a un accent peu intense ; si donc *i* et *u* inaccentués se sont amuis en arménien ancien, tandis que *a*, *e* et *o* inaccentués se maintenaient, c'est sans doute que, comme en grec septentrional, les voyelles accentuées étaient sensiblement plus longues que les inaccentuées. De même il n'y a aucune raison de croire que le roman initial (dit latin vulgaire) ait eu un accent bien intense, et l'on sait que l'accent du français actuel comporte une très faible intensité : il est donc arbitraire de supposer que, à un moment de son histoire, le français aurait eu un accent du type allemand ou russe ; ce n'est pas un accent peu intense de ce type qui aurait entraîné les grandes réductions de voyelles observées dans le passage du latin en français : si *camera* est devenu *chambre*, ce n'est pas parce que *ca-* était accentué avec beaucoup plus de force que *me* et *ra*. Le développement est plutôt du type grec septentrional ; et, en effet, il a atteint la voyelle la plus ouverte, *a*, moins que les autres voyelles : *nouu*(*m*) a donné *neuf*, tandis que *noua*(*m*) a donné *neuve* ; dès avant les premiers textes français, l'*u* de *nouu*(*m*) s'est amui, tandis que l'*a* de *noua*(*m*) s'est réduit à *e* dit muet ; et ce n'est qu'au cours de l'histoire du français, et même à date relativement récente, que l'*e* de *neuve* a cessé de se prononcer.

La recherche des conditions générales où les changements phonétiques peuvent se produire n'est qu'à ses débuts. Mais, dès maintenant, on y rencontre beaucoup de faits certains. Par exemple, une consonne appuyée, c'est-à-dire suivant une autre consonne, est plus forte, c'est-à-dire plus susceptible de durer, qu'une consonne placée entre voyelles ; le *-t-* intervocalique de *dōtāre* a disparu dans fr. *douer* ; mais le *-t-* précédé de *-k-* de *lac-*

tūca a subsisté dans fr. *laitue*. Il y a, dans ces constatations de quoi guider le linguiste.

Toutefois des formules exactes ne se laissent pas encore trouver dans tous les cas. Il faut du moins toujours observer les conditions propres à certains phonèmes. Par exemple le *y*, loin d'entrer dans la règle des consonnes appuyées, tend à se fondre dans la prononciation d'une consonne précédente, si bien qu'il perd souvent son existence propre en altérant profondément cette consonne. Mais on pourrait écrire — Schleicher l'a déjà tenté, prématurément — une étude d'ensemble sur les combinaisons de *y* avec les consonnes précédentes et les phénomènes multiples qui en résultent.

Il y a des cas complexes : des groupes de consonnes intervocaliques comme *-mn-* admettent des traitements variés. En italien, *-n-* est traitée comme si elle était appuyée, et *somnu(n)* aboutit à *sonno* ; en français au contraire *-mn-* forme un groupe qui appartient tout entier à l'explosion, et c'est *m* qui y domine, d'où *somme* ; en espagnol, enfin, il se produit un phonème complexe, et l'aboutissement est *sueño*, où l'on ne retrouve proprement ni *m* ni *n*. A essayer de ramener ces traitements à une formule une on perdrait sa peine. Mais il ne devient pas impossible pour cela de poser des formules : les formules ne s'appliquent pas à l'ensemble du développement entre un moment initial et divers aboutissements plus ou moins lointains. Les formules de la phonétique générale évolutive s'appliquent seulement au moment où a lieu chaque changement partiel. Si *somnu(m)* a donné *sonno* en italien, c'est que, au moment où a eu lieu l'assimilation, *-mn-* formait un groupe de deux consonnes distinctes où *-n-* était traitée comme appuyée. Si *somnu(m)* a donné *somme* en français, c'est que, au moment du changement, *-mn-* était étroitement groupé et que *-n-* avait perdu son autonomie. Il sort de là la conclusion que, dans un groupe tel que *-mn-*, la nasale *-n-* est sujette à perdre son caractère de consonne appuyée, ce qui tient au caractère spécial des nasales *m* et *n*.

Enfin — et ceci est le trait essentiel —, les formules de la phonétique générale évolutive indiquent des possibilités, non des nécessités. On peut déterminer comment une consonne placée entre voyelles est sujette à s'altérer; mais ceci n'implique pas qu'elle s'altérera. Un *-k-* placé entre voyelles est sujet à s'altérer soit en spirante gutturale *x* (le *ch* allemand) soit en occlusive sonore *g* ; et le *x* et le *g* peuvent subir ultérieurement d'autres altérations dues à la position intervocalique ; par exemple le *k* roman a fini par passer à *i* en français dans un cas tel que *auca* donnant *oie* (maintenant prononcé *wa*) ; mais un *k* intervocalique de l'indo-européen a subsisté partout en slave, et le *k* qu'on prononce aujourd'hui dans polonais *piekę* « je cuis » existe dans cette première personne depuis l'indo-européen (l'*i* de fr. *cuis* représente le même élément — après bien des transformations successives).

En matière de morphologie, les choses se passent de manière moins rigide qu'en matière de phonétique. La morphologie est le domaine de la survivance.

Telle catégorie grammaticale qui s'explique par la mentalité des hommes d'une certaine société persiste longtemps après que cette mentalité a disparu. Les hommes qui voient des forces internes derrière les phénomènes naturels et les événements peuvent classer les notions en animées et inanimées et, dans l'animé, distinguer systématiquement ce qui est mâle ou assimilé au mâle de ce qui est femelle ou assimilé à une femelle. On conçoit alors comment l'arbre qui produit des fruits peut être nommé au féminin et le fruit au « neutre » (genre inanimé) ; mais pareille mentalité n'existait plus dans la Rome du Ier siècle ap. J.-C., et une opposition comme celle entre *pirus* « poirier » (féminin) et *pirum* « poire » n'était plus qu'une survivance. Dans le français actuel, il serait très choquant de dire *la ciel* ou *le terre* ; mais personne ne saurait dire pourquoi, au point de vue d'un Français d'aujourd'hui ; ce n'est qu'en vertu d'une tradition séculaire que les substantifs s'opposent ainsi.

Telle catégorie qui a eu un sens fort et un large emploi tend à perdre de son sens et beaucoup de son usage. Dans le verbe latin, le « subjonctif » était une pièce essentielle, et le rôle du subjonctif domine la syntaxe latine. Mais le rôle du subjonctif n'a cessé de se restreindre depuis l'époque romane commune. Aujourd'hui, en français, le subjonctif sert de substitut à des formes manquantes de l'impératif : à côté de *viens, venez*, on a *qu'il vienne, qu'ils viennent*. Et il y a des tours de phrases où le subjonctif est encore de rigueur : *je veux qu'il vienne, il faut qu'il vienne*. Mais le français actuel n'a plus le sens de la valeur du subjonctif. De là vient que la langue populaire en restreint l'usage de plus en plus : le français littéraire a *je doute qu'il vienne* ; mais en français courant, on tend à dire : *je doute s'il viendra*. Un tour tel que : *s'il fait cela et qu'il dise ce qu'il pense, il aura tort* est maintenant archaïque.

Sous le bénéfice de cette réserve, la morphologie évolue comme la phonétique, d'après des formules générales. Un fait fondamental de l'évolution des langues indo-européennes en fournira un exemple.

En indo-européen commun, le nom et le verbe étaient également des formes fléchies. Il n'y avait pas une forme du nom, celle du cas sujet par exemple, dont les autres auraient été des dérivées : un nom se composait d'un ensemble de formes casuelles dont aucune ne commandait les autres ; il n'y a pas en latin un mot *loup*, mais un ensemble *lupus, lupe, lupum, lupī, lupō*, où aucune de ces cinq formes ne sert à former les autres, pas plus qu'elle ne sert à former le pluriel *lupī, lupīs, lupōrum, lupīs* ; il n'y a pas un nom de l' « homme », mais un ensemble *homō, hominem, hominis, hominī, homine,* et ainsi toujours. De même il n'y a pas une forme de « aller » qui commande les autres, mais *eō, īs, it,* et de *ībam, ībō, eam, īrem,* etc.

Or, au cours du développement des langues indo-européennes, les formes nominales ont, plus ou moins tôt, mais partout, tendu à réduire leur flexion : sauf les formes anciennes de l'indo-iranien, aucune langue indo-européenne ne présente au complet toutes les formes de la déclinaison. Et, avec le temps, la déclinaison a disparu dans une grande partie des langues, ainsi dans les langues romanes, en anglais, en persan ; ou bien la forme du cas sujet a pris une importance dominante, ainsi en slave, ou a servi de point de départ aux autres formes, ainsi en arménien. Au contraire, toutes ces langues ont conservé une flexion verbale, sinon aussi riche que la flexion indo-européenne, du moins variée, complexe, pourvue de formes multiples. La comparaison entre le verbe français, encore si encombré de formes diverses, et le nom français, invariable, est significative.

Ce n'est pas un accident : le sémitique commun avait une flexion nominale à trois cas distingués par des désinences; seuls, l'akkadien le plus ancien (babylonien) et l'arabe classique en donnent une idée; l'akkadien récent et l'arabe moderne l'ont éliminée; au contraire, le verbe sémitique a partout, et même dans une langue aussi évoluée que l'araméen, une conjugaison riche de formes. Et, d'une manière générale, on observe souvent ce contraste entre l'absence de flexion nominale et la richesse de la flexion verbale.

Ceci tient à la nature du nom et du verbe. Le nom indique une notion de caractère permanent : une chose, une personne, ou une qualité; le verbe indique un procès, qu'il s'agisse d'une action proprement dite ou d'un état : « il mange, il dort », etc. Par cela même qu'il indique une notion permanente, le nom comporte une forme unique ou, tout au plus, une forme principale dont se tirent les autres. Au contraire, le verbe, qui indique un procès, comporte des formes personnelles et l'expression de nuances qui sont diverses suivant les langues, mais qui peuvent être nombreuses. Le fait que les formes du nom et celles du verbe obéissent à des tendances divergentes résulte donc de la nature des choses.

Qu'il s'agisse de phonétique ou de morphologie, les formules générales du développement s'expliquent par des conditions communes à tous les hommes, ou, du moins, à tout un type de civilisation.

IX.

LES INNOVATIONS SPECIFIQUES.

Les conditions communes à tous les hommes ou à tout un type de civilisation ont pour effet de maintenir l'unité. Mais il y a aussi des conditions propres à certains groupes d'hommes, à certains temps, à certaines régions. De ces conditions spéciales, il résulte qu'une langue une à un moment donné se différencie avec le temps en parlers divers.

Tant qu'il n'intervient pas de tendances caractéristiques nouvelles, le système linguistique ne change pas essentiellement. Les modifications qu'amènent les tendances générales au changement ne provoquent pas d'ordinaire la création de phonèmes nouveaux. Ainsi, quand un phonème perd par dissimilation l'un de ses éléments, la langue ne garde pas le phonème singulier qui résulterait purement et simplement de la dissimilation. Si, dans un ancien *veneno,* la première des deux *n* perd l'abaissement du voile du palais, le résultat n'est pas une occlusive très faible sans nasalisation, comme on l'attendrait ; ce phonème dont l'articulation serait trop peu marquée a été aussitôt remplacé par un phonème de type voisin existant dans la langue, et l'on a *veleno* ; l'*n* dénasalisée n'a eu qu'une existence virtuelle. Le changement une fois accompli, il y a eu dans la langue une *l* dans un mot de plus, mais non pas un type phonique de plus.

On observe fréquemment des faits de ce genre. A date relativement récente, le français a perdu *l* mouillée qui était un phonème fréquent ; ce n'a été qu'une simplification de l'articulation, et il n'est intervenu aucun grand fait qui ait pu provoquer une variation notable du système articulatoire. Aussi l'élimination de *l* mouillée n'a-t-elle pas eu pour conséquence de créer quelque phonème nouveau : *l* mouillée s'est simplifiée en *y*, c'est-à-dire en un phonème qui se trouvait couramment, dans *miel* et dans *pied*, dans *yeux* et dans *aïeux*, etc. Après l'innovation, le français a eu beaucoup plus de *y* qu'il n'en avait précédemment ; mais le système phonique n'a pas été modifié.

Les anciens **sw*- et **tw*- ont abouti en arménien à *k*ʿ (*k* dit aspiré). Ceci a supprimé des *w* consonnes ; mais le *k*ʿ ainsi produit est identique à des *k*ʿ d'autre origine qui existaient en arménien, dans des mots comme *k*ʿ*an* « que » ou *lk*ʿ*anem* « je laisse ».

En morphologie, le simple jeu de l'analogie ne fait qu'élargir l'emploi de catégories existantes sans en créer de nouvelles. Si en français *dites, faites* sont remplacés par *disez, faisez*, la langue n'a rien de nouveau ; car on dit *lisez, taisez*, etc. Si, en grec, on a fait *hestēka* en face de *hestamen*, avec -*k*- d'après *ethēka, ethemen*, ce n'a été que pour faciliter la formation du parfait ; on a beaucoup élargi l'emploi d'un procédé qui se trouvait à l'origine dans deux ou trois formes seulement ; mais il n'a été constitué ni un procédé nouveau ni une catégorie nouvelle.

Or, l'expérience montre qu'il se crée des procédés nouveaux et phoniques et morphologiques. Dans toutes les langues indo-européennes, on observe des innovations qui ont transformé la langue en y introduisant des phonèmes jusque-là inusités, des formes grammaticales neuves et même des catégories neuves. Il y a donc des innovations spécifiques qui ne se laissent ramener à aucune tendance de caractère universel.

L'indo-européen avait des occlusives sourdes, articulées avec force, et où l'explosion n'était suivie d'aucun souffle, de sorte que les vibrations vocaliques des voyelles commençaient aussitôt après l'explosion : c'est le type des sourdes *p, t, k* telles qu'on les rencontre dans les langues romanes, en baltique, en slave, en iranien, par exemple. Il y avait des occlusives sonores, plus faiblement articulées et accompagnées de vibrations du larynx dès le moment de l'implosion : tels sont les *b, d, g* des mêmes langues. En arménien, en germanique et même en celtique (ici moins radicalement), ces types ont été remplacés par d'autres : l'articulation de *p, t, k* est devenue plus faible, et un souffle s'est inséré entre l'explosion et le début des vibrations laryngales qui caractérisent la voyelle ; quant à *b, d, g,* les vibrations laryngales n'ont pas commencé au moment même de l'implosion, de sorte que les *b, d, g* ainsi prononcés se sont rapprochés du type sourd (en Français des *b, d, g* de cette sorte font l'effet de *p, t, k* mal réussis). Le système de l'articulation des occlusives a donc changé tout entier. Le système nouveau était moins stable que l'ancien ; et, tandis que les langues qui sont demeurées fidèles au type indo-européen présentent encore, surtout à l'initiale des mots ou en position appuyée (après consonne), l'état indo-européen, il y a eu des innovations de détail nombreuses qui ont dérangé souvent le parallélisme — ainsi l'arménien et le celtique ont perdu l'ancien *p* initial. — Ce grand changement du type articulatoire a des chances de s'être produit brusquement, lorsque l'indo-européen s'est étendu à des populations nouvelles.

Mais il y a d'autres changements spécifiques qui se produisent après le changement de langue, souvent longtemps après, et qui n'ont lieu que peu à peu, sans doute en vertu de tendances héréditaires acquises. Ainsi, dans les langues germaniques, la prononciation de la voyelle de la syllabe suivante se prépare au moment où s'émet une voyelle; ceci a eu de grandes conséquences. Chacune des voyelles indo-européennes était ou nettement prépalatale ou nettement postpalatale ; au contraire, le germanique, tendant à préparer un *i* ou un *y* quand il prononce un *a* ou un *o* par exemple, a créé des voyelles mixtes, du type de *eu* et *ü* du français. Ces voyelles diffèrent beaucoup d'une langue germanique à l'autre, d'un parler germanique à un autre. Il en est résulté des complications extrêmes, notamment en vieil anglais et dans les parlers nordiques. Ces innovations se sont achevées, pour une large part, peu avant la période historique de ces langues ou encore pendant la période historique. Mais elles étaient amorcées depuis longtemps. Ici encore, il y a une innovation spécifique du germanique, et cette innovation atteint l'ensemble du système vocalique qu'elle transforme radicalement.

Des innovations appartenant à une même tendance peuvent aboutir à des dates assez différentes. Ainsi les formes consonantiques correspondant aux voyelles *i* et *u*, à savoir *y* et *w*, se sont éliminées en grec. Mais, alors que *y* était inconnu de tous les parlers grecs dès une époque bien antérieure à l'époque historique et que, lors de l'emprunt de l'alphabet phénicien, le signe du yod a été affecté à noter seulement la voyelle *i*, au contraire *w* (qu'on nomme en grec digamma) a subsisté longtemps à l'époque historique ; certains textes de beaucoup de parlers ont encore tous les anciens *w* ; d'autres, qui gardent le *w* initial, ont perdu déjà le *w* intervocalique ; quelques particularités de l'attique montrent que, avant l'époque historique, le *w* y a existé. L'élimination de *w* n'a donc eu lieu que longtemps après l'élimination de *y*. Néanmoins, les deux phénomènes résultent d'une même tendance fondamentale; or, la langue doit à ce changement une part notable de la différence d'aspect qu'on remarque entre l'indo-européen et le grec.

Quelquefois l'unité de la tendance est masquée par des accidents particuliers. Ainsi, en roman, à l'initiale du mot, les anciens *y* et *w* ont tendu à prendre un caractère plus consonantique que celui qu'avaient *i* et *u* consonnes en latin. L'*i* consonne est devenu *dz̧*, d'où fr. *j* ; ainsi de *iam* on a en italien *gia* et en français *ja-* (dans *jamais*). L'*u* consonne est devenu aussi plus consonne ; mais le latin avait déjà *f* labio-dentale ; la sonore correspondante, *v*, manquait ; la forme consonantique de *u* est venue remplir cette case vide du système phonétique. Or, comme *f* et *v* sont des consonnes stables, ces spirantes ont subsisté. Le traitement de *i* consonne a cessé ainsi d'être parallèle à celui de *u* consonne. — A l'intérieur du mot, entre voyelles, le latin avait perdu *i* consonne bien avant l'époque historique alors que *u* consonne subsistait ; le *u* consonne intérieur a passé aussi à la prononciation spirante, soit *v*.

En matière de morphologie, il y a de même des tendances spécifiques, et il en résulte à la fois des pertes et des innovations.

Il y a plusieurs langues indo-européennes où s'est perdue la distinction des genres grammaticaux. Sans doute cette distinction ne répond plus à rien dans la mentalité des hommes civilisés depuis des siècles. Mais elle survit dans les langues romanes, en allemand, en slave, etc. où, à défaut de valeur sémantique, elle tient une place importante parmi les procédés grammaticaux. On ne voit pas qu'elle avait plus de raisons de disparaître en arménien — où la distinction des thèmes en *-o-* et des thèmes en *-ō-* s'est maintenue dans la déclinaison des substantifs — qu'en allemand par exemple. Or, l'arménien n'en a aucune trace. Il y a vraiment ici tendance spécifique ; car, dans l'iranien occidental, qui est voisin, la distinction des genres s'est également abolie. Et il n'est pas fortuit que, ni en élamite, ni dans le caucasique du Sud, qui sont les langues voisines, il n'y ait pas non plus de différences de genre. L'action du substrat se manifeste ici par la disparition d'une catégorie grammaticale.

Malgré la richesse de sa flexion nominale, le letto-lituanien a perdu la distinction du masculin et du neutre, alors que les formes qui servaient à exprimer le masculin et le neutre demeuraient nettement distinctes. Les finales *-as* et *-ą* qui devaient en lituanien répondre à lat. *-us* et *-um* étaient claires. Or, dès le letto-lituanien, la distinction du masculin et du neutre — encore existante en vieux prussien à la fin du XIVᵉ siècle — est abolie. Il y a ici une tendance spécifique, et il est permis de l'attribuer à des mélanges de populations de langue finnoise avec celles qui parlaient le letto-lituanien.

Plus caractéristiques que les pertes, sont les innovations. En indo-européen, les thèmes verbaux exprimaient l' « aspect » : le duratif (présent), le procès pur et simple (aoriste), le procès achevé (parfait), le procès arrivant à son terme (aspect déterminé), etc. Ils n'exprimaient jamais le temps. Le futur n'avait pas d'expression propre. La distinction du présent et du passé s'exprimait par des différences de désinences qui servaient aussi à d'autres usages, ou par l'augment, qui est un mot accessoire, et qui se rencontre seulement dans un groupe dialectal : indo-iranien, arménien et grec. Or, le celtique et l'italique ont largement développé l'expression du temps par les thèmes verbaux, et le temps est devenu, dans la conjugaison celtique ou italique, un élément dominant. Rien de plus spécifique. Le celtique et l'italique ne sont pas pour cela des langues « mixtes » ; car le temps s'exprime par des procédés nouveaux qui diffèrent du celtique à l'italique, de l'osco-ombrien au latin, et même de l'osque à l'ombrien.

Partout ou presque en indo-européen, les formes casuelles cessent de suffire à indiquer des relations concrètes : pour marquer le lieu où l'on est, le lieu où l'on va, on ajoute des prépositions aux cas, en attendant que les prépositions se substituent tout à fait aux formes casuelles. Or, non seulement le lituanien a maintenu un locatif fortement caractérisé et qui se suffit, mais il a constitué un illatif distinct de l'accusatif. Il est vrai que locatif et illatif comprennent des postpositions ; mais, au point de vue lituanien, les formes de ces cas ont le caractère de formes casuelles pures et simples. Ici encore, la concordance avec le type finnois est frappante. Mais pas une forme grammaticale finnoise n'a pour cela pénétré en lituanien pas plus qu'en lette.

Pour comprendre le développement des langues, il importe de marquer l'opposition des tendances spécifiques aux formules générales. Innovations et conservations prennent ainsi une valeur significative que, autrement, elles n'auraient pas.

Il va de soi qu'il ne suffit pas de reconnaître les innovations spécifiques pour déterminer le « substrat » sur lequel une langue nouvellement adaptée s'est développée, ni même pour donner le droit d'affirmer l'existence d'un substrat étranger. Mais on a ainsi le moyen de poser la question, et une première amorce pour la résoudre.

Quoi qu'on pense du « substrat », il faut, dans tous les cas, distinguer des changements normaux les changements qui caractérisent une langue. Quand on considère des périodes étendues de l'évolution, il ressort presque toujours quelques innovations caractéristiques.

X.

NÉCESSITÉ DE PRÉCISIONS NOUVELLES.

Les conclusions énumérées jusqu'ici sont de types divers. Les unes portent sur des faits historiques : l'italien, l'espagnol, le portugais, le français, le roumain sont des formes divergentes prises par le latin. Les autres expriment des possibilités générales : une consonne placée entre voyelles est sujette à des altérations auxquelles échappe une consonne initiale ou une consonne appuyée sur une consonne précédente. Mais, dans les deux cas, le procédé de démonstration est le même. On dira : les concordances entre italien, espagnol, etc. sont telles qu'elles excluent le hasard ; les différences de traitement entre les consonnes intervocaliques et les consonnes initiales ou appuyées sont telles qu'elles excluent le hasard. Tout le problème de méthode porte donc sur l'appréciation du hasard : le linguiste, observant que les dés retombent souvent du même côté, conclut qu'ils sont pipés. Mais combien de concordances et quelles concordances faut-il rencontrer pour qu'on considère comme inadmissible le caractère fortuit de la rencontre ?

On ne saurait faire intervenir le calcul des probabilités. Car les concordances ne sont pas de même sorte ni de même valeur. Le calcul des probabilités ne s'applique utilement qu'à des faits qualitativement identiques et comportant une commune mesure. Tel n'est pas le cas des faits linguistiques.

Par cela même que la qualité, les détails particuliers en doivent être appréciés en chaque cas, il dépend du tact, du jugement, du bon sens des linguistes de tirer des faits le parti qui convient. Tous les linguistes ne sentent pas les choses de la même manière. Tel qui a un sens juste des transformations phonétiques se rend mal compte de la façon dont les sujets parlants se comportent vis-à-vis d'un mot donné ; les dons et les connaissances qui permettent de bien comprendre le développement des formes grammaticales ne sont pas ceux grâce auxquels on devine les actions capricieuses de l' « étymologie populaire ». Ce « coefficient personnel » explique beaucoup de divergences entre les linguistes. Et il y a des esprits faux qui ne sont capables d'apprécier la juste valeur d'aucun rapprochement. Mais, avec de l'exercice et en examinant des séries de faits bien établis, on acquiert une expérience grâce à laquelle on peut juger correctement des divers types de faits comparés. Voici des exemples de rapprochements délicats, où l'on apercevra les difficultés de la preuve.

On constate que, en brittonique comme en gaélique, les consonnes intervocaliques sont traitées autrement que les consonnes initiales. Ce qui tend à indiquer qu'il y a là une tendance commune, c'est que le traitement intervocalique s'applique, dans les deux groupes, aux consonnes initiales de mots étroitement liés à des mots précédents dans la phrase. Mais le résultat diffère d'un groupe à l'autre : en gaélique, les occlusives sourdes intervocaliques deviennent des spirantes sourdes : *t* passe à _ɸ_ et _k_ à _x_ ; en brittonique, au contraire, ces occlusives deviennent des sonores : _t_ passe à _d_, et _k_ à _g_. Le changement ne s'applique pas aux sonores géminées qui sont plus fortement articulées que les simples. — La différence de traitement entre les sourdes gaéliques et les sourdes brittoniques doit tenir à des propriétés des deux langues considérées ; car, en gaélique, certaines consonnes sonores par nature tendent à s'assourdir, ce qui n'arrive pas en brittonique. En gaélique, le _w_ initial a passé à _f_, tandis que, en brittonique, il a donné _g_^{v} ; en gaélique, le _y_ initial a passé à _h_ qui s'est amuï, tandis que au moins _y_ subsistait en brittonique. — Le linguiste observe la coïncidence. Qui admet que la forme des changements phonétiques tient avant tout à la structure phonique de la langue tiendra la coïncidence pour significative. Mais tout le monde ne souscrira peut-être pas à cette conclusion.

Les concordances de faits particuliers ont d'autant plus de valeur probante que ces faits répondent moins à des tendances générales. Pour établir que les langues italiques proviennent d'une même langue commune (d'origine indo-européenne) que les langues celtiques, on dispose de quelques coïncidences. L'une des plus curieuses est que l'ancienne forme du nom de nombre « cinq », qui est _*penk_^{w}_e_ (_pañca_ en sanskrit, _pénte_ en grec), est passée à italo-celt. _*k_^{w}_enk_^{w}_e_, comme on le voit par _coic_ de l'irlandais et _quinque_ du latin. La valeur probante de la concordance est difficile à déterminer. On sait que le fait n'est pas particulier à ce nom de nombre ; le même phénomène se retrouve partout où l'on rencontre _p_ et _k_^{w} dans deux syllabes consécutives en italique et en celtique. — L'innovation ne résulte manifestement pas d'une tendance spécifique de l’italo-celtique. Les assimilations de ce genre relèvent de tendances générales à l’assimilation dans deux syllabes consécutives (type de **seš*uras* devenant *šešuras* en lituanien). Mais ce qu’il faudrait savoir, c’est quelle est la puissance de cette tendance ; or, on l’ignore. On a peu d’occasions d’observer la consécution *p*... k^w... — Il est certain qu’il y avait, en indo-européen occidental, parenté entre k^w et *p* et que partout, régulièrement en grec (avec des limitations définies), en osco-ombrien et en celtique, sporadiquement en germanique, k^w est passé à *p*. Le passage de **p*... k^w... à *k^w... k^w... en italo-celtique est l’un des faits qui expriment cette parenté. Mais l’assimilation ne s’est pas produite en grec ; l’italique et le celtique ont donc ici une innovation qui les oppose au grec. En germanique, il y a eu une assimilation, mais en sens inverse, et au lieu de *k^w... k^w... il y a eu **p*... *p*... représenté par *f-f* dans un exemple isolé, soit *fimf* en gotique. On voit par là que la tendance à l’assimilation était naturelle, mais que la forme de l’assimilation comportait deux possibilités. Ce qui caractérise l’italo-celtique, c’est, d’une part, le sens dans lequel s’est produite l’assimilation, de l’autre, la constance du phénomène. La concordance est probante, mais moins qu’il ne semble au premier abord.

Par bonheur, le comparatiste a souvent le moyen de trouver des démonstrations relativement objectives.

Le type le plus caractéristique du problème qui se pose au comparatiste historien est l’établissement d’une étymologie par voie de comparaison. Pour donner la preuve d’un rapprochement, il faut montrer que certaines concordances observées ne peuvent être fortuites. Le fait que le mot *erku* « deux » de l'arménien est à rapprocher de grec *dyo, dyō,* latin *duo,* etc., comme on l'a vu p. 6, fournit un bon exemple du type de preuves qu'on peut employer. Au premier abord, le rapprochement surprend, et l'on hésite à l'accepter. En réalité, il est sûr, et l'on peut le montrer aisément.

Tout d'abord, la série des noms de nombre arméniens, de « un » à « dix », est indo-européenne. La forme *erku* de « deux » est la seule qui ne soit pas transparente du premier coup. Un emprunt, peu vraisemblable *a priori* pour un nom de nombre dans une ancienne langue indo-européenne où le vieux fonds de vocabulaire est si bien conservé dans l'ensemble, est d'autant moins admissible qu'aucune des langues connues près de l'arménien ne fournit une forme pareille pour « deux » et que par suite on ne voit pas ici comment l'arménien aurait pris *erku* au dehors.

En second lieu, les noms de nombre « trois » et « quatre », qui avaient en indo-européen la flexion du pluriel, ont en arménien au nominatif la forme du nominatif avec -*k*ʻ final, tandis que, à partir de « cinq », qui est le premier des noms de nombre indo-européens non fléchis, il n'y a pas -*k*ʻ : *erek*ʻ « trois », *čork*ʻ « quatre », mais *hing* « cinq », *vec* « six », etc. Or, le nom de nombre « deux » avait en indo-européen une flexion, qui naturellement était celle du duel. A l'époque historique, l'arménien n'avait plus le duel, et sans doute depuis longtemps ; la forme *erku* ne peut donc s'expliquer à date historique. Et il est remarquable que *erku* soit demeuré sans signe du pluriel, alors que la catégorie du duel était abolie en arménien. Cette survivance en arménien d'une trace du duel indo-européen est frappante ; elle rappelle ce que l'on observe dans lat. *duo*, gr. *dyō̆*, etc.

Le traitement **erk-* de **dw-* initial est étrange. Mais on a pu montrer qu'il se retrouve dans tous les cas où figure ce même groupe initial ; deux autres rapprochements, aussi bons pour le sens que celui de *erku* avec lat. *duo*, l'établissent. Et le fait s'explique en arménien (sur tout cela, voir p. 31).

Du reste, si *erku* était un ancien dissyllabe et si l'*e* y avait existé originairement, l'-*u* final n'aurait pu se maintenir ; car, dans tout dissyllabe dont l'arménien a hérité, la voyelle de la syllabe finale s'est amuie. L'-*u* de *erku*, qui répond si naturellement à l'ancien *-*ō* attesté par hom. *dyō*, v. sl. *dŭva*, etc., n'est demeuré que parce que, à la date où se sont amuies les voyelles des syllabes finales, le mot n'était pas dissyllabique.

Toutefois l'étrangeté du traitement phonétique pourrait laisser des doutes à certains esprits. La démonstration est acquise dès que l'on montre les variations concordantes de la forme de « deux » en arménien et dans les autres langues indo-européennes. Comme un *ō* indo-européen est représenté par *u* en arménien, la forme *erku* repose sur **duwō*, conservé par exemple dans la forme homérique *dyō* et dans v. sl. *dŭva*, comme dans védique *d(u)vā́*. Mais il y avait une autre forme **duwo*, avec *o* bref final, qui est la forme usuelle en grec, soit *dyo*. Or, à côté de *erku* « deux », l'arménien a *erko-tasan* « douze » (deux et dix), où le type à *o* bref est maintenu. Et ce n'est pas tout : au premier terme des composés, l'indo-européen avait **dwi-*, et non **dwo-*, ainsi *dvi-* en sanskrit, *di-* en grec, *bi-* en latin ; or, l'arménien a *erki-* dans *erkeam* « de deux ans ». Cette triple série de variations concomitantes exclut le hasard. La démonstration est ainsi achevée.

Il n'est pas toujours possible de trouver des concordances aussi complètes et aussi nombreuses. Il résulte de là que le degré de probabilité des étymologies varie d'un cas à l'autre. Dans quelques cas, comme celui qui vient d'être cité, la preuve ne laisse rien à désirer, et elle atteint un degré de rigueur qui se rencontre rarement en matière historique. Dans d'autres, le rapprochement n'est que possible, et parfois il est malaisé d'en montrer même la vraisemblance. Entre les deux extrêmes, il y a tous les degrés de probabilité : l'un des défauts les plus graves de beaucoup de dictionnaires étymologiques, c'est qu'ils n'indiquent pas assez les différences de valeur entre les rapprochements. Si le linguiste ne dispose pas de coefficients pour noter ces différences, il ne manque cependant pas de moyens de les faire sentir et de les signaler.

Pour faire progresser la linguistique historique, il importe de préciser, de systématiser et d'étendre les recherches. Car les théories reposent sur des données incomplètes, vagues, livrées par le hasard plutôt que choisies.

Il faut des observations toujours plus précises : à chaque fois qu'on a observé les données de plus près, on a pu obtenir des résultats nouveaux. Pour les langues actuellement parlées, on est loin d'avoir mis en œuvre tous les moyens dont on dispose. La phonétique a besoin de laboratoires et, dans ces laboratoires, de linguistes exercés à manier les instruments. Et il ne suffit pas d'observer les faits tels qu'ils sont à un moment donné. Un article récent du regretté Poirot, paru dans les *Mélanges Andler*, a montré comment on peut saisir un changement au moment même où il se réalise, et en voir le mécanisme. Pour les langues anciennes, le linguiste doit recourir à une philologie de précision : on s'est parfois imaginé que le linguiste peut se contenter d'à peu près philologiques ; il a besoin tout au contraire de tout ce que les méthodes philologiques les plus exactes permettent de précision et de rigueur.

Il faut instituer des recherches systématiques. Presque rien n'a été fait jusqu'ici pour établir par des recherches méthodiques les théories linguistiques. Seuls, quelques psychologues ont amorcé des séries d'expériences dont la plupart n'ont pas été poursuivies par d'autres.

Il importerait par exemple d'étudier avec soin en quelles mesures les habitudes linguistiques acquises se transmettent de génération en génération. On ne sait pas si un enfant apprend plus facilement la langue de ses parents qu'il n'apprendrait une langue différente, et surtout une langue de type différent. Il y a là toute une enquête à faire, et qui n'offre pas de difficultés particulières.

Certains faits donneraient lieu de croire que les enfants de parents connaissant bien plusieurs langues ou des enfants de parents bilingues seraient plus aptes à bien apprendre eux-mêmes des langues diverses que des enfants de parents parlant l'un et l'autre une seule langue et la même.

Le fait que des innovations semblables semblent apparaître vers le même temps chez les enfants nés dans les mêmes conditions relèverait de l'hérédité des habitudes acquises.

On entrevoit ici tout un ensemble d'observations et même de recherches expérimentales qui jetteraient sur le développement linguistique une lumière nouvelle.

Du reste, l'expérience des linguistes est trop étroite. Si les grandes langues sont décrites souvent avec minutie, il s'en faut de beaucoup que les parlers locaux le soient avec assez de précision. En dehors des grandes langues, on n'a que peu de descriptions exactes ; or, il est nécessaire de voir comment fonctionnent des langues de type aussi différent qu'il est possible, et dans des conditions aussi variées qu'il est possible.

Chose plus grave, il n'y a presque pas d'état de langue qui ait été observé, décrit d'une manière exacte, précise, complète. Ce que donnent les descriptions phonétiques, les grammaires, les dictionnaires, ce sont ou des cas typiques, plus ou moins arbitrairement choisis, ou des normes. Mais il n'y a guère de domaine de l'ensemble duquel on puisse dire quel est en fait l'état linguistique. Or, ce qui intéresse le linguiste, ce ne sont pas les normes, c'est la façon dont la langue est mise en usage.

Rien ne diffère plus d'un état de langue qu'un autre état de langue. Autre chose est le parler d'un village rural où tout le monde est à peu près au même niveau social, où tout le monde a à peu près la même culture, autre chose est le parler d'une ville où il y a des hommes de conditions diverses, d'éducations diverses, de cultures diverses. Autre chose est une ville où tout le monde emploie la même langue, autre chose une ville, comme Constantinople, où se juxtaposent des populations parlant cinq ou six langues distinctes. Les conditions varient suivant les mœurs : dans telle ville du Levant, où les hommes savent plus ou moins quatre ou cinq langues et emploient au dehors ces langues suivant les cas, les femmes confinées à la maison peuvent n'avoir qu'une seule langue. La variété des situations suivant les gens est infinie.

On croit savoir ce que c'est que le français. En réalité personne ne sait vraiment comment parle l'ensemble des sujets ni dans un village français, ni dans une ville de province française, ni moins encore à Paris. Qui voudrait savoir comment évoluera le français devrait chercher en quelle mesure s'emploient les parlers locaux à la campagne et dans les villes, quelle forme spéciale affecte le français dans chaque province, dans chaque ville et, dans chaque ville, chez les gens de chaque condition sociale, de chaque profession, de chaque groupe.

Le français des grammaires et des dictionnaires est connu ; mais ce n'est qu'un ensemble de prescriptions. Ce qui importe au linguiste, c'est de connaître comment les gens qui parlent français se comportent vis-à-vis des règles. Or, on n'a là-dessus que des idées vagues ; il n'a été fait aucune enquête méthodique, à peine quelques sondages partiels.

Et pourtant le fait que le français commun devient en France la langue de tout le monde alors que, au début du XIX^e siècle, il était la langue d'une très petite minorité d'hommes cultivés, a radicalement changé les conditions du développement de la langue et ne peut manquer d'y déterminer des innovations considérables.

Ce qui est vrai du français l'est de toutes les grandes langues du monde.

L'histoire constate que des langues communes s'étendent. Mais, en ce qui concerne le passé, la façon dont a eu lieu cette extension n'est pas déterminable. Et, quant au présent qu'on pourrait observer, on le laisse s'écouler sans presque le regarder. Avant 1914, le russe devenait à Tiflis la langue commune. Mais on n'a jamais indiqué par quels procédés, par quels intermédiaires. Le besoin d'avoir une langue commune dans une ville où vivaient côte à côte des gens dont la langue maternelle était le géorgien, l'arménien, le turc, le russe et bien d'autres idiomes encore, poussait à l'emploi du russe pour lequel travaillaient l'école et l'administration et qui avait le prestige d'apporter la civilisation. Mais tel menu trait de mœurs pouvait y contribuer largement : les femmes arméniennes ne se plaçaient pas comme servantes, et les familles arméniennes aisées avaient à leur service des femmes russes, notamment pour élever les enfants ; le russe devenait ainsi, pour nombre de petits Arméniens, surtout dans les familles les plus considérées, les plus influentes, une langue maternelle.

Le berbère, ancienne langue du pays, et l'arabe, langue des envahisseurs qui ont apporté l'islamisme, coexistent dans les territoires coloniaux de l'Afrique du Nord administrés par la France. L'arabe, qui est la langue de la religion et une grande langue littéraire, tend, au moins en Algérie, à éliminer le berbère, qui n'est qu'un ensemble de parlers locaux. En revanche, les populations islamiques, fidèles à leur religion et à leurs traditions, n'acceptent de la civilisation européenne que certains de ses éléments matériels ; elles n'adoptent donc pas le français en dehors de certaines nécessités extérieures. Les progrès de la civilisation qu'amène une administration européenne se trouvent dès lors pousser à l'abandon du berbère et à la généralisation de l'arabe, si bien que la domination française travaille sans le vouloir consciemment à étendre le domaine de l'arabe.

Il faudrait savoir, en cas de bilinguisme, comment chaque langue agit sur l'autre et quelles réactions résultent de là.

L'Amérique fournit des exemples remarquables, qu'il faudrait étudier en détail, et que les linguistes ont à peine effleurés.

L'Amérique du Nord a été colonisée par des colons en majorité anglais aux XVII^e^ et XVIII^e^ siècles. Le pays actuel des États-Unis n'a cessé qu'à la fin du XVIII^e^ siècle d'être une colonie anglaise, et le Canada est, aujourd'hui, un *Dominion* de l'Empire britannique ; des Anglais ont émigré chaque année pour les États-Unis au cours du XIX^e^ et du XX^e^ siècle. La langue littéraire n'a jamais cessé d'être la même des deux côtés de l'Atlantique ; les normes de la grammaire sont les mêmes ; l'école donne le même enseignement grammatical ; il y a même, surtout dans l'Est des États-Unis, un purisme anglais chez les gens les plus cultivés. — D'autre part, l'influence des indigènes n'existe pas. Les uns ont été éliminés, les autres ont été parqués dans des régions particulières ; il n'y a pas eu mélange des blancs avec les anciens habitants du pays. — Et pourtant, cent cinquante ans seulement après la déclaration de l'indépendance, l'anglais parlé en Amérique diffère notablement de celui qui s'emploie en Angleterre. Un critique littéraire distingué, M. Mencken, va jusqu'à parler — un peu tendancieusement — de deux langues distinctes.

La divergence s'explique.

Tout d'abord, l'anglais est moins résistant au changement que le francais. La grammaire est moins rigide ; l'anglais est surtout caractérisé par des expressions idiomatiques qui sont plus sujettes à se modifier que ne le sont des formes grammaticales proprement dites. Le purisme est d'ailleurs moins puissant en Angleterre qu'en France, moins organisé, moins populaire.

Les États-Unis sont devenus une nation puissante et qui ne sent, vis-à-vis de l'ancienne patrie, aucune dépendance, à qui une dépendance, même limitée à la langue, pèserait. Si des groupes intellectuels restent liés à la vieille culture anglaise, surtout dans la région d'ancienne colonisation, dans l'Est, c'est le fait de lettrés dont l'influence sur l'ensemble du pays n'est pas dominante. L'école ne dépend pas du gouvernement central et ne saurait jouer le rôle fortement unificateur qu'elle joue en France par exemple.

Et surtout les États-Unis sont un pays d'immigration. Il y est allé des émigrants de toutes les régions de l'Europe : irlandais pour qui, même s'ils parlaient anglais, l'anglais était le résultat d'une acquisition peu ancienne, allemands, scandinaves, lettes et lituaniens, polonais, russes, finlandais, roumains, juifs de l'Europe orientale, italiens, grecs, arméniens, et même syriens. L'anglais est la langue commune des États-Unis ; mais cette langue commune est parlée par beaucoup de sujets dont la langue maternelle était autre ou qui, même s'ils sont nés aux États-Unis, ont entendu de leurs parents une autre langue. Dans la plupart des régions, ces émigrants nouveaux sont beaucoup plus nombreux que les descendants d'anciens colons. L'anglais est donc parlé aux États-Unis par une majorité de gens chez qui la tradition de l'anglais est récente, et dont beaucoup sont encore bilingues. Au fur et à mesure que grandit l'importance du centre et de l'ouest des États-Unis où la proportion des immigrés récents est plus forte que dans l'est et où il y a moins de traditions britanniques, la langue risque plus de se corrompre.

En somme, peu de forces conservatrices, beaucoup de forces qui tendent au changement. Le fait que, dès maintenant, l'anglais évolue de manière différente en Angleterre et aux États-Unis n'a rien que de naturel.

Il serait singulièrement instructif de suivre les destinées de l'anglais dans le monde.

En se proposant d'examiner un domaine complexe tel que celui des langues du Caucase, l'Institut pour l'étude comparative des civilisations a donc le sentiment juste des besoins les plus urgents de la linguistique. Ce sont des enquêtes de cette sorte qui fourniront des renseignements précis et complets sur des états de langue, et qui, par là, permettront de renouveler les doctrines. On attend avec confiance les résultats que, par les soins d'observateurs fins et exercés, de linguistes profondément avertis, donnera pareille enquête, et l'on souhaite que ces résultats provoquent à bref délai des enquêtes pareilles sur d'autres domaines. De ces enquêtes et de la rigueur avec laquelle elles seront menées dépend l'avenir de la linguistique.

TABLE DES MATIÈRES

Pages.

Avant-propos. V

I. Définition de la méthode comparative. 1

II. Les langues communes. 13

III. Des preuves employées. 22

IV. Du développement linguistique entre l'époque de communauté et l'époque historique. 43

V. Les dialectes. 53

VI. La géographie linguistique. 60

VII. La notion de langue mixte. 72

VIII. Les formules générales de changement. 84

IX. Les innovations spécifiques. 95

X. Nécessité de précisions nouvelles. 103

ACHEVE D'IMPRIMER
LE 14 JUIN 1965
PAR JOSEPH FLOCH
MAITRE - IMPRIMEUR
A MAYENNE
n°2368

EN VENTE A LA MÊME LIBRAIRIE

A. MEILLET et A. VAILLANT, Professeurs au Collège de France et à l'Ecole des Hautes Etudes. — **GRAMMAIRE SERBO-CROATE.** In-8°, 304 pages.

BEAULIEUX. — **Observations sur l'Orthographe de la langue française.** Transcriptions, commentaire du manuscrit de Mézeray (1673).

BEAULIEUX. — **Histoire de l'Orthographe française.** 2 volumes in-8° de 366 et 132 pages. Couronné par l'Académie. Prix Saintour. Tome I : Formation de l'orthographe. Tome II : Les accents et autres signes auxiliaires.

BOURGUET (E). — **Le Dialecte laconien.**

CAHEN. — **Etude sur le Vocabulaire religieux du vieux scandinave. La Libation.** In-8°.

CARRA DE VAUX. — **La Langue étrusque, sa place parmi les langues.** In-8°.

DAUZAT. — **Le Français populaire de Saint-George-de-Didonne.**

ETIENNE. — **La Langue française depuis les origines jusqu'à la fin du XI^e siècle.** Tome I : Phonétique, déclinaison, conjugaison.

LANGLARD. — **La Liaison dans le français.**

MEILLET. — **Linguistique historique et linguistique générale.** Tome I.

MEILLET. — **Les Dialectes Indo-Européens.** In-8°.

— **Mélanges de Philologie et d'Histoire.** Offerts à M. Antoine Thomas, par ses élèves et ses amis.

MISTRAL. — **La Génési.** Traducho en prouvençau. Emé lou latin de la vulgato vis-à-vis e lou francé en dessouto per J.-J. Brousson E., en testo, lou retra dou Felibré.

RENOU. — **La Valeur du parfait dans les hymnes védiques.**

RINGENSON. — **Etude sur la palatalisation de K devant une voyelle antérieure en français.** In-8°.

SECHEHAYE. — **Essai sur la structure logique de la phrase.**

TANQUEREY. — **L'Evolution du verbe en anglo-français (XII^e et XIV^e siècles).** In-8°.

SJOESTEDT-JONVAL. — **Description d'un parler irlandais de Kerry.**

SUETONI TRANQUILLI. — **De grammaticis et rhetoribus edidit apparatu et critico instruxit.** R. P. Robinson, 1926, in-8° raisin de 80 p.

THOMAS. — **Mélanges d'étymologie française,** première série, in-8° de 304 p. (collection linguistique t. 22).

THUROT. — **De la prononciation française depuis le commencement du XVI^e siècle.** D'après le témoignage des grammairiens, t. I. in-8°. IV-568 p.

T. I seul réimprimé de cet ouvrage très rare.